Introduction générale

La tâche la plus difficile de ce cours est de définir le mot "statistique"

Citations humoristes :

- la statistique est la première des sciences inexactes.
- c'est l'art de mentir mathématiquement.
- je ne crois aux statistiques que lorsque je les ais moi-même falsifiées.

Définition

La statistique est un ensemble de techniques mathématiques permettant de recueillir, décrire et interpréter des données statistiques, pour pouvoir en déduire des prévisions et des décisions.

La statistique comporte deux aspects :

La Statistique descriptive : C'est l'ensemble des méthodes de collecte et traitement des données statistiques relatives à un échantillon, sans chercher à établir des prévisions pour une population plus grande.

La description des données peut se faire par :

- Une présentation sous forme detableaux.
- Une représentation graphique.
- Un résumé numérique par le calcul de certaines grandeurs typiques (moyenne, variance, ...).
- Une étude des éventuelles corrélations entre variables.

La statistique inférentielle : C'est l'ensemble des méthodes qui permettent à partir de l'étude d'un échantillon d'induire des informations sur une population plus grande.

- Elle nécessite un choix judicieux de l'échantillon : il doit être représentatif de la population (cf. exemple des élections américaines de 1936).
- Elle utilise des modèles théoriques de référence : les lois de probabilités.

La statistique trouve ses applications dans plusieurs domaines scientifiques de différentes natures : économie, biologie, chimie, physique, sociologie, médecine, industrie, ...

Objectif de ce cours :

- Comprendre les principales techniques de la statistique **descriptive** à une dimension et à deux dimensions.
- Être capable de mettre en oeuvre ces techniques de manière appropriée dans un contexte donné.
- En résumé, Au bout de ce cours, l'étudiant doit être capable de répondre aux questions suivantes (Voir TD) :

Exercice 1

La température dans une région A est relevée chaque jour, à la même heure, pendant une année. Les résultats obtenus sont rassemblés dans le tableau suivant :

Température	20	22	23	25	26	27	29	30	32
Nombre de jours	10	15	30	40	55	70	60	50	35

1. Préciser la population étudiée, le caractère étudié et sa nature.
2. Tracer les diagrammes qui permettent de représenter cette distribution statistique.
3. Déterminer les paramètres de tendance centrale et les paramètres de dispersion des températures dans la région A.

Exercice 2

La distribution conjointe d'un groupe d'étudiants selon la note obtenue en module de statistique (caractère X) et le nombre d'absence aux cours et travaux dirigés (caractère Y) est représentée dans le tableau suivant :

X \ Y	0	1	2	3	4
$[0, 6[$	0	4	3	8	12
$[6, 10[$	3	7	10	12	18
$[10, 12[$	6	8	9	8	7
$[12, 14[$	7	6	5	3	2
$[14, 16[$	8	3	2	0	0
$[16, 20[$	9	4	1	0	0

1. Compléter ce tableau avec les distributions marginales des deux caractères X et Y.
2. Tracer l'histogramme et calculer le mode de la distribution marginale de X.
3. Tracer la courbe cumulative et calculer la médiane de la distribution marginale de X.
4. Quelle est la proportion des étudiants :

   a) ayant validé le module de statistique (note supérieure à 10) ?
   b) n'ayant pas validé le module de statistique (note comprise entre 0 et 7) ?
   c) ayant obtenu un rattrapage (note comprise entre 7 et 10) ?
5. Calculer la moyenne et la variance de la distribution marginale de X.

1. Calculer la moyenne et la variance de la distribution marginale de Y.
2. Calculer les moyennes conditionnelles de X pour chaque valeur de Y. Conclure.
3. Calculer la covariance des deux variables X et Y.
4. Déterminer la droite de régression de X en fonction de Y.
5. Calculer le coefficient de corrélation linéaire. Commenter.

Plan du cours

1. Statistique descriptive à une dimension (séries simples)
   - I- Définitions
   - II- Organisation des données
     - II.1- Tableaux statistiques
     - II.2- Représentations graphiques
   - III- Réduction des données
     - III.1- Paramètres de position
     - III.2- Paramètres de dispersion
2. Statistique descriptive à 2 dimensions (séries doubles)
   - I- Définitions
   - II- Paramètres d’une série double
   - III- Ajustement

UNIVERSITÉ HASSAN II DE CASABLANCA
FACULTÉ DES SCIENCES ET TECHNIQUES

STATISTIQUE DESCRIPTIVE

Pr. Abdelhak FAHSI

Département de Mathematiques

Introduction générale

La tâche la plus difficile de ce cours est de définir le mot ”statistique”

Citations humoristes :
- la statistique est la première des sciences inexactes.
- c’est l’art de mentir mathématiquement.
- je ne crois aux statistiques que lorsque je les ais moi-même falsifiées.

Définition

La statistique est un ensemble de techniques mathématiques permettant de recueillir, décrire et interpréter des données statistiques, pour pouvoir en déduire des prévisions et des décisions.

La statistique comporte deux aspects :

La Statistique descriptive : C'est l'ensemble des méthodes de collecte et traitement des données statistiques relatives à un échantillon, sans chercher à établir des prévisions pour une population plus grande.

La description des données peut se faire par :

- Une présentation sous forme detableaux.
- Une représentation graphique.
- Un résumé numérique par le calcul de certaines grandeurs typiques (moyenne, variance, ...).
- Une étude des éventuelles corrélations entre variables.

La statistique inférentielle : C'est l'ensemble des méthodes qui permettent à partir de l'étude d'un échantillon d'induire des informations sur une population plus grande.

- Elle nécessite un choix judicieux de l'échantillon : il doit être représentatif de la population (cf. exemple des élections américaines de 1936).
- Elle utilise des modèles théoriques de référence : les lois de probabilités.

La statistique trouve ses applications dans plusieurs domaines scientifiques de différentes natures : économie, biologie, chimie, physique, sociologie, médecine, industrie, ...

Objectif de ce cours :

- Comprendre les principales techniques de la statistique **descriptive** à une dimension et à deux dimensions.
- Être capable de mettre en oeuvre ces techniques de manière appropriée dans un contexte donné.
- En résumé, Au bout de ce cours, l'étudiant doit être capable de répondre aux questions suivantes (Voir TD) :

Exercice 1

La température dans une région A est relevée chaque jour, à la même heure, pendant une année. Les résultats obtenus sont rassemblés dans le tableau suivant :

Température	20	22	23	25	26	27	29	30	32
Nombre de jours	10	15	30	40	55	70	60	50	35

1. Préciser la population étudiée, le caractère étudié et sa nature.
2. Tracer les diagrammes qui permettent de représenter cette distribution statistique.
3. Déterminer les paramètres de tendance centrale et les paramètres de dispersion des températures dans la région A.

Exercice 2

La distribution conjointe d'un groupe d'étudiants selon la note obtenue en module de statistique (caractère X) et le nombre d'absence aux cours et travaux dirigés (caractère Y) est représentée dans le tableau suivant :

X \ Y	0	1	2	3	4
$[0, 6[$	0	4	3	8	12
$[6, 10[$	3	7	10	12	18
$[10, 12[$	6	8	9	8	7
$[12, 14[$	7	6	5	3	2
$[14, 16[$	8	3	2	0	0
$[16, 20[$	9	4	1	0	0

1. Compléter ce tableau avec les distributions marginales des deux caractères X et Y.
2. Tracer l'histogramme et calculer le mode de la distribution marginale de X.
3. Tracer la courbe cumulative et calculer la médiane de la distribution marginale de X.
4. Quelle est la proportion des étudiants :

   a) ayant validé le module de statistique (note supérieure à 10) ?
   b) n'ayant pas validé le module de statistique (note comprise entre 0 et 7) ?
   c) ayant obtenu un rattrapage (note comprise entre 7 et 10) ?
5. Calculer la moyenne et la variance de la distribution marginale de X.

1. Calculer la moyenne et la variance de la distribution marginale de Y.
2. Calculer les moyennes conditionnelles de X pour chaque valeur de Y. Conclure.
3. Calculer la covariance des deux variables X et Y.
4. Déterminer la droite de régression de X en fonction de Y.
5. Calculer le coefficient de corrélation linéaire. Commenter.

Plan du cours

1. Statistique descriptive à une dimension (séries simples)
   - I- Définitions
   - II- Organisation des données
     - II.1- Tableaux statistiques
     - II.2- Représentations graphiques
   - III- Réduction des données
     - III.1- Paramètres de position
     - III.2- Paramètres de dispersion
2. Statistique descriptive à 2 dimensions (séries doubles)
   - I- Définitions
   - II- Paramètres d'une série double
   - III- Ajustement

Chapitre 1

Statistique descriptive à une dimension

(Séries simples)

I- Définitions

1. Population (notée P) : Ensemble des individus (personnes ou objets) sur lesquels on veut effectuer l'étude statistique.

Exemple

- Les étudiants d'une université
- Les véhicules automobiles vendues au Maroc en 2007
- Les accidents sur l'autoroute au mois d'août.

2. Individu : Tout élément (notée ω) de l'ensemble P s'appelle un individu.

3. Echantillon : Groupe resreint d'individus prélevés dans la population définie au préalable (sous-ensemble de P).

On étudie un échantillon lorsque la population est impossible (ou très difficile) à étudier dans son intégralité.

4. Caractère statistique (noté X, Y, ...) : Caractéristique relative à chacun des individus de la population, que l'on décide d'observer et analyser.

Exemple

- Le poids des enfants âgés de 5 ans dans une région,
- Le salaire mensuel des employés d'une entreprise,
- La couleur des voitures vendues au Maroc en 2010,

Dans une même étude statistique, on pourra considérer plusieurs caractères simultanément sur une même population.

On distingue deux types de caractères :

- Caractère qualitatif : C'est un caractère non mesurable. Les valeurs du caractère ne sont pas numériques.

  Exemple
  La couleur, la nationalité, la situation familiale, ...

- Caractère quantitatif : C'est un caractère mesurable à valeurs dans $\mathbb{R}$.

  Exemple
  Le poids, la taille, le salaire, ...

Remarque :

Dans la suite de ce cours, on n'étudiera que les caractères quantitatifs, les autres peuvent s'y ramener par un codage.

Ainsi, dans la suite, un caractère X désigne tout simplement une application:

$$X : P \to \mathbb{R}$$

5. Modalités (notées x_i) : toute valeur prise par le caractère X s'appelle modalité de X.

L'ensemble des modalités d'un caractère X (valeurs distinctes prises par X) est noté :

$$X(P) = \{x_1, x_2, \ldots, x_k\}$$

k désigne le nombre de modalités de X.

Exemple

- Le caractère "$X =$ puissance fiscale (en CV)" peut avoir comme modalités : 6, 7, 8, 9, 10, 11 et plus.

6. Effectif (noté n_i) : L'effectif d'une modalité x_i est le nombre d'individus présentant cette modalité, ($n_i = card(X^{-1}\{x_i\})$).

L'effectif total, noté N, est le nombre total des individus de la population (appelé aussi taille de la population). On a :

$$\sum_{i=1}^{k} n_i = N$$

7. Effectif cumulé (noté N_i) : On suppose les modalités ordonnées :

$$x_1 < x_2 < \ldots < x_k$$

L'effectif cumulé d'une modalité x_i est le nombre d'individus de la population pour lesquels la valeur du caractère est inférieure ou égale à x_i, ($N_i = card(X^{-1}\{x_1, x_2, \ldots, x_i\})$. On a :

$$N_i = \sum_{j=1}^{i} n_j = N_{i-1} + n_i$$

8. Fréquence ou proportion (notée f_i) : La fréquence d'une modalité x_i, est le rapport $f_i = \dfrac{n_i}{N}$.

C'est la proportion des individus de la population présentant cette modalité. On a :

$$\sum_{i=1}^{k} f_i = 1$$

Remarque

La fréquence f_i appartient à l'intervalle $[0, 1]$.

Parfois, on note les fréquences en pourcentage (avec le symbole %) en les multipliant par 100.

9. Fréquence cumulée (notée F_i) : La fréquence cumulée d'une modalité x_i est la proportion d'individus pour lesquels la valeur du caractère est inférieure ou égale à x_i :

$$F_i = \frac{N_i}{N} = \sum_{j=1}^{i} f_j = F_{i-1} + f_i$$

Remarque

- Lorsque le nombre des modalités d'un caractère est élevé (supérieur à 15), on est généralement conduit à regrouper les modalités en classes de la forme :

$$[a_i; a_{i+1}[\ \ (ou\ ]a_i; a_{i+1}]), \quad i = 1, 2, \ldots, k$$

- Les classes $[a_i; a_{i+1}[$ peuvent avoir une même amplitude ou des amplitudes différentes.

Exemple

Le caractère "Revenus mensuels (en dh) de 100 employés d'une entreprise" peut avoir comme modalités groupées : $]4000; 5000],]5000; 6000],]6000; 8000]$, plus de 8000.

On appelle effectif d'une classe $]a_i; a_{i+1}]$ (noté n_i) le nombre d'individus ayant une modalité appartenant à cette classe, $(n_i = card(X^{-1}(]a_i; a_{i+1}]))$.

On appelle effectif cumulé de la classe $]a_i; a_{i+1}]$, (noté N_i) le nombre d'individus pour lesquels la valeur du caractè est inférieure ou égale à a_{i+1},
(N_i=effectif des i premières classes).

On a des définitions analogues pour les fréquences et fréquences cumulées.

Règles de construction des classes

- Fixer un nombre de classes ni trop petit ni trop grand (généralement de 5 à 15).
- Choisir des bornes qui, autant que possible, permettront des calculs simples.
- Considérer des classes adjacentes fermées d'un côté et ouvertes de l'autre.

10. Type d'un caractère - Série statistique :

Caractère discret :

Si les modalités d'un caractère X sont considérées sous forme de valeurs en nombre fini $\{x_1, x_2, \ldots, x_k\}$, on dit que X est un caractère de type discret.

La famille (x_i, n_i) ou (x_i, f_i) est appelée série statistique discrète associée à X.

10. Type d'un caractère - Série statistique :

Caractère continu :

Si les modalités d'un caractère X sont groupées sous forme de classes $(]a_1, a_2], \ldots,]a_k, a_{k+1}])$, on dit que X est un caractère de type continu.

La famille $(]a_i, a_{i+1}], n_i)$ est appelée série statistique continue associée à X.

II- Organisation des données

II.1- Tableaux statistiques

Un tableau statistique consiste à résumer et présenter les données d'une série statistique associée à un caractère X sous forme de distributions d'effectifs ou de fréquences en fonction des modalités :

Tableau statistique :

Modalités x_i	Effectifs n_i	...
x_1	n_1	
x_2	n_2	
$\vdots$	$\vdots$	
x_k	n_k	
Total	N	

Tableau statistique :

Modalités x_i	Fréquences f_i	...
x_1	f_1	
x_2	f_2	
$\vdots$	$\vdots$	
x_k	f_k	
Total	1	

Exemple1 : Dans un échantillon composé de 50 familles, on enquête sur le nombre d'enfants par famille.
Les résultats de l'enquête statistique sont :

1	0	5	2	2	1	2	1	2	4
4	7	1	3	2	5	4	6	3	1
1	6	1	3	8	1	3	5	2	3
3	0	3	4	6	4	1	7	2	0
2	0	1	2	2	3	2	5	6	2

Questions : - Population étudiée ?
- Caractère étudié ?
- Modalités ?
- Nature du caractère ?

Tableau statistique :

Nombre d'enfants	Effectif n_i	...
0	4	
1	10	
2	12	
3	8	
4	5	
5	4	
6	4	
7	2	
8	1	
Total	50	

Tableau statistique détaillé :

Nombre d'enfants	Effectif n_i	Eff. cum. N_i	Fréquence f_i	Fréq. cum. F_i	...
0	4	4	0.08	0.08	
1	10	14	0.2	0.28	
2	12	26	0.24	0.52	
3	8	34	0.16	0.68	
4	5	39	0.1	0.78	
5	4	43	0.08	0.86	
6	4	47	0.08	0.94	
7	2	49	0.04	0.98	
8	1	50	0.02	1	
Total	50		1		

Exemple2 : En mesurant la taille de 50 étudiants de la FST, on a obtenu les résultats suivants (en cm) :

152	151,5	160	165	170
159	168	161	164	156
158.5	167	157	170,5	161,5
169	156	158,5	160,5	152
156.5	166	152,5	170	165
154	170	165	155.5	166,5
162,5	152,5	168	169	158
157	161	154,5	162	158
153,5	157,5	163	155	153
160	169,5	154	161	162

- Population étudiée ?
- Caractère étudié ?
- Modalités : Le nombre des modalités observées étant élevé, il est donc nécessaire de les grouper en classes.
- Nature du caractère ?

On peut considérer les classes suivantes :

$[151; 155[$
$[155; 159[$
$[159; 163[$
$[163; 167[$
$[167; 171[$

Tableau statistique :

Taille (en cm)	Effectif n_i	...
$[151; 155[$	10	
$[155; 159[$	12	
$[159; 163[$	11	
$[163; 167[$	7	
$[167; 171[$	10	
Total	50	

Tableau statistique :

Taille (en cm)	Effectif n_i	Eff. cum. N_i	Fréquence f_i	Fréq. cum. F_i	...
$[151; 155[$	10	10	0.2	0.2	
$[155; 159[$	12	22	0.24	0.44	
$[159; 163[$	11	33	0.22	0.66	
$[163; 167[$	7	40	0.14	0.8	
$[167; 171[$	10	50	0.2	1	
Total	50		1		

II.2- Representations graphiques

- Bien qu'un tableau statistique résume toute l'information d'une distribution statistique, la représentation graphique permet de visualiser et de déceler les principales caractéristiques de la distribution statistique (tendance, symétrie, dispersion, concentration, . . .).

- Bien qu'un tableau statistique résume toute l'information d'une distribution statistique, la représentation graphique permet de visualiser et de déceler les principales caractéristiques de la distribution statistique (tendance, symétrie, dispersion, concentration, . . .).

- La représentation graphique des données relatives à un caractère repose sur la proportionnalité des longueurs (caractère discret) ou des surfaces (caractère continu) aux effectifs des différentes modalités du caractère.

- Ainsi, suivant le type du caractère étudié, on utilise différents modes de représentations graphiques.

1. Caractère quantitatif discret :

a. Diagramme en bâtons :

Chaque modalité est représentée par un trait vertical (bâton) dont la hauteur est égale à l'effectif (ou à la fréquence) de la modalité.

1. Caractère quantitatif discret :

Exemple1 : Nombre d'enfants par famille

Nombre d'enfants	Effectif n_i	Eff. cum. N_i	Fréquence f_i	Fréq. cum. F_i	...
0	4	4	0.08	0.08	
1	10	14	0.2	0.28	
2	12	26	0.24	0.52	
3	8	34	0.16	0.68	
4	5	39	0.1	0.78	
5	4	43	0.08	0.86	
6	4	47	0.08	0.94	
7	2	49	0.04	0.98	
8	1	50	0.02	1	
Total	50		1		

1. Caractère quantitatif discret :

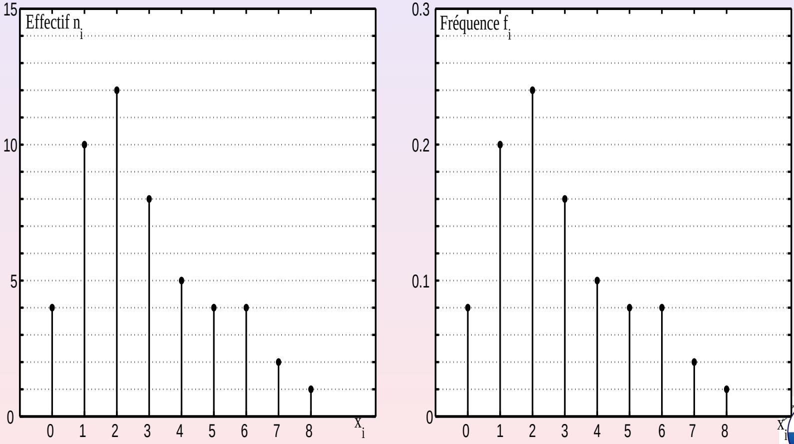


Figure: Diagramme en bâtons

1. Caractère quantitatif discret :

b. Polygone des effectifs :

Le polygone des effectifs (ou des frèquences) est construit en joignant par des segments de droites les sommets des bâtons du diagramme en bâtons.

1. Caractère quantitatif discret :

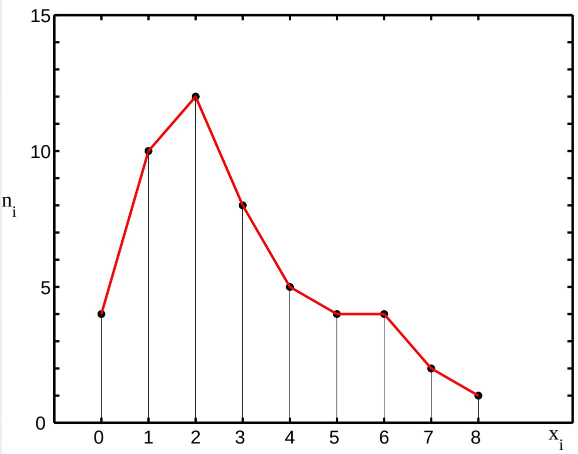


Figure: Polygone des effectifs.

1. Caractère quantitatif discret :

c. Courbe cumulative - Fonction de répartition :

Soient $x_1 < x_2 < \ldots < x_k$ les modalités ordonnées du caractère.

À partir des effectifs cumulées N_i, on définit la fonction de répartition des effectifs $G(x)$, définie de $\mathbb{R}$ vers $[0, N]$, par :

$$
\begin{array}{lll}
G(x) = 0, & \text{si} & x < x_1 \\
G(x) = N_i, & \text{si} & x_i \leq x < x_{i+1} \quad (1 \leq i \leq k-1) \\
G(x) = N, & \text{si} & x \geq x_k
\end{array}
$$

$G(x)$ représente le nombre d'individus ayant une modalité inférieure ou égale à x.

1. Caractère quantitatif discret :

La courbe cumulative (ou la courbe des effectifs cumulés) est la représentation graphique de la fonction de répartition $G(x)$.

1. Caractère quantitatif discret :

Exemple1 : Nombre d'enfants par famille

Nombre d'enfants	Effectif n_i	Eff. cum. N_i
0	4	4
1	10	14
2	12	26
3	8	34
4	5	39
5	4	43
6	4	47
7	2	49
8	1	50
Total	50	

1. Caractère quantitatif discret :

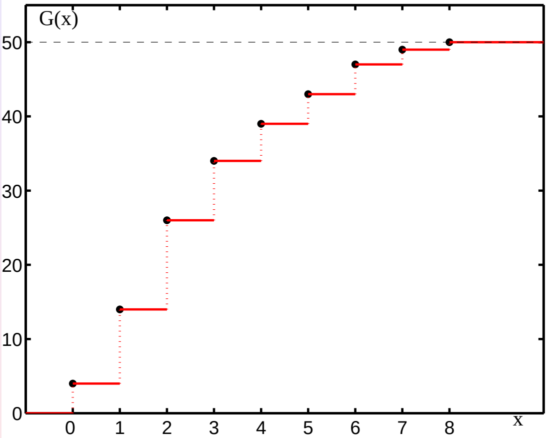


Figure: Courbe cumulative.

1. Caractère quantitatif discret :

Remarque :

À partir des fréquences cumulées F_i, on définit la fonction de répartition $F(x)$, définie de $\mathbb{R}$ vers $[0, 1]$, par :

$$\begin{array}{lll} F(x) = 0, & \text{si} & x < x_1 \\ F(x) = F_i, & \text{si} & x_i \leq x < x_{i+1} \quad (1 \leq i \leq k-1) \\ F(x) = 1, & \text{si} & x \geq x_k \end{array}$$

$F(x)$ représente la proportion d'individus ayant une modalité inférieure ou égale à x. On a :

$$F(x) = \frac{G(x)}{N}$$

1. Caractère quantitatif discret :

Remarque :

La courbe cumulative est aussi la représentation graphique de la fonction de répartition $F(x)$.

Elle s'appelle aussi la courbe des fréquences cumulées.

1. Caractère quantitatif discret :

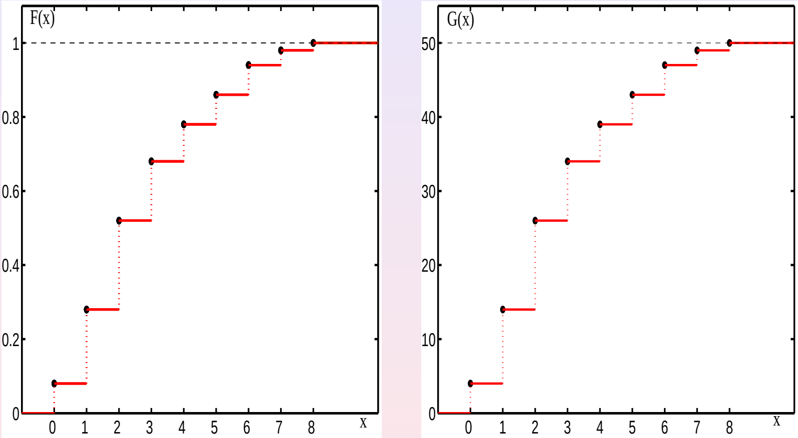


Figure: Courbe cumulative.

2. Caractère quantitatif continu :

a. Histogramme - Polygone des effectifs :

Les modalités sont présentées sous forme de classes $[x_i; x_{i+1}[$, $1 \leq i \leq k$.

Pour tracer l'histogramme, chaque modalité $[x_i; x_{i+1}[$ doit être représentée par un rectangle dont la base a_i est égale à l'amplitude de la classe ($a_i = x_{i+1} - x_i$) et dont la hauteur h_i est telle que la surface $S_i = a_i \times h_i$ du rectangle soit proportionnelle à l'effectif n_i (ou à la fréquence f_i) de la classe :

$$S_i = C \times n_i$$

C est une constante arbitraire à choisir.

2. Caractère quantitatif continu :

Remarque :
On distingue deux cas :

1er cas : Toutes les classes ont une même amplitude a.

Dans ce cas, on représente chaque modalité par un rectangle dont la hauteur est égale à l'effectif (ou à la fréquence) de la modalité. La base de tous les rectangles étant la même (égale à l'amplitude a).

Ceci revient à choisir la constante C égale à l'amplitude a.

2. Caractère quantitatif continu :

Exemple2 : Taille des étudiants

Taille (en cm)	Effectif n_i
$[151; 155[$	10
$[155; 159[$	12
$[159; 163[$	11
$[163; 167[$	7
$[167; 171[$	10
Total	50

2. Caractère quantitatif continu :

Histogramme : Toutes les classes ont une même amplitude.

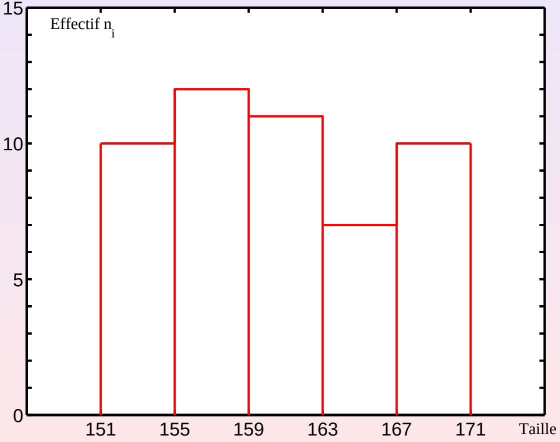


Figure: Histogramme

2. Caractère quantitatif continu :

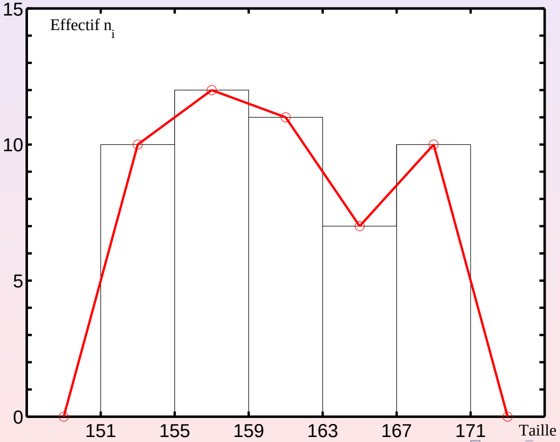
Polygone des effectifs :

Lorsque toutes les classes ont une même amplitude, le polygone des effectifs est construit en joignant par des segments de droites les milieux des côtés supérieurs des rectangles dans l'histogramme.

Les extrémités rejoignent l'axe des abscisses.

2. Caractère quantitatif continu :

Polygone des effectifs : Toutes les classes ont une même amplitude *a*.



2. Caractère quantitatif continu :

2ème cas : Les classes n'ont pas toutes la même amplitude.

Dans ce cas, pour construire l'histogramme dans, chaque modalité $[x_i; x_{i+1}[$ doit être représentée par un rectangle dont la base est égale à l'amplitude a_i de la classe $[x_i; x_{i+1}[$ et dont la hauteur h_i est proportionnelle à la densité d'effectif $d_i = \dfrac{n_i}{a_i}$ correspondante :

$$h_i = C \times \frac{n_i}{a_i}$$

C est une constante de proportionnalité à choisir.

2. Caractère quantitatif continu :

h_i est appelée effectif corrigé (ou fréquence corrigée) de la modalité $[x_i; x_{i+1}[$.

Le choix de la constante C est arbitraire (on peut prendre $C = 1$ par exemple).

Cependant, pour simplifier les calculs et/ou pour tracer le polygone des effectifs, il faut choisir C égale au plus grand diviseur commun des amplitudes a_i.

2. Caractère quantitatif continu :

Exemple : Taille des étudiants (les classes n'ont pas toutes la même amplitude)

Taille (en cm)	amplitude a_i	Effectif n_i	Eff. corrigé h_i
$[151; 155[$	4	10	10
$[155; 159[$	4	12	12
$[159; 167[$	8	18	9
$[167; 171[$	4	10	10
Total		50	

$$h_i = C \times \frac{n_i}{a_i} \quad \text{avec } C = PGDC(a_i) = 4.$$

2. Caractère quantitatif continu :

2ème cas : Histogramme avec effectif corrigé.

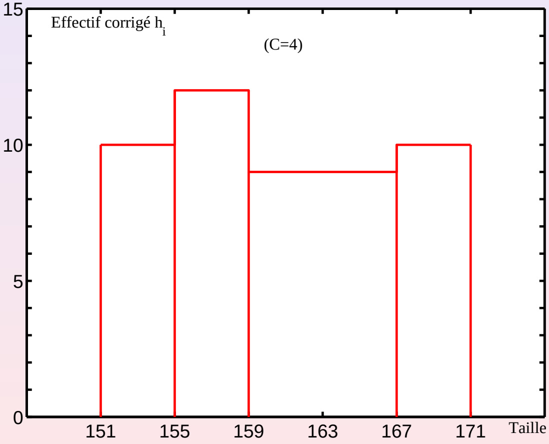
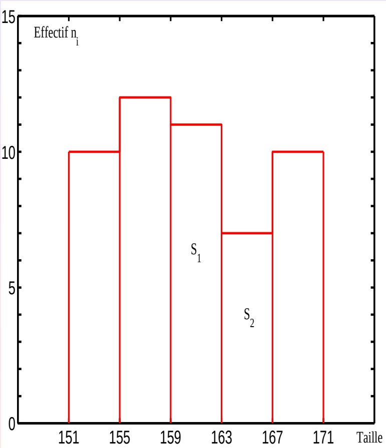
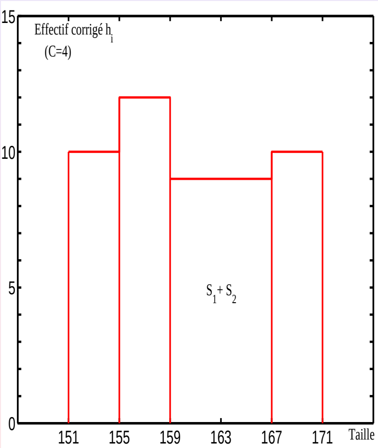


Figure: Histogramme

2. Caractère quantitatif continu :





2. Caractère quantitatif continu :

Polygone des effectifs : les classes n'ont pas la même amplitude.

On choisit $C = PGCD(a_i)$. Donc chaque a_i est un multiple de C, càd $a_i = k_i \times C$.

Pour construire le polygone des fréquences dans le cas où les classes n'ont pas la même amplitude, on considère chaque rectangle de base $a_i = k_i \times C$ comme réunion de k_i sous-rectangles de même base C.

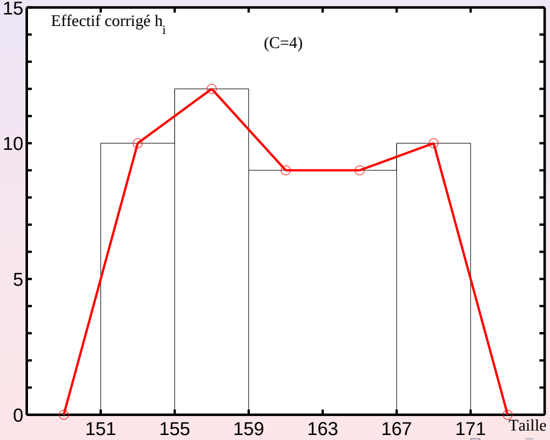
2. Caractère quantitatif continu :

Le polygone des effectifs est construit en joignant successivement, par des segments de droites, les milieux des côtés supérieurs des sous-rectangles de base C dans l'histogramme.

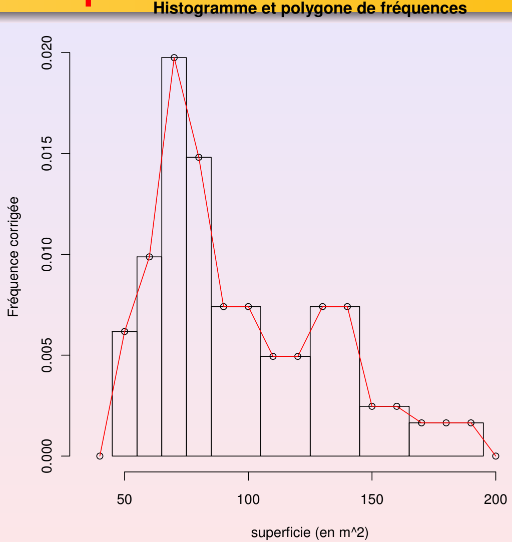
Les extrémités du polygone rejoignent l'axe des abscisses.

2. Caractère quantitatif continu :

Polygone des effectifs : Les classes n'ont pas la même amplitude.



2. Caractère quantitatif continu :



2. Caractère quantitatif continu :

Remarque

L'aire des rectangles de l'histogramme est égale à l'aire délimité par le polygone des effectifs.

2. Caractère quantitatif continu

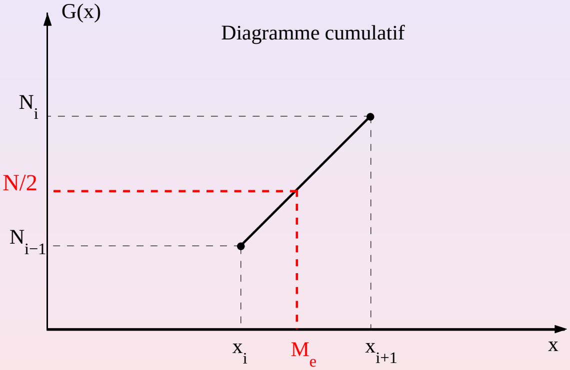
b. Courbe cumulative - Fonction de répartition :

Remarques

- Comme pour le cas discret, la fonction de répartition des effectifs $G(x)$, définie de $\mathbf{R}$ vers $[0, N]$, représente le nombre des individus de la population pour lesquels la valeur du caractère est inférieure ou égale à x.
- La fonction de répartition $F(x)$, définie de $\mathbf{R}$ vers $[0, 1]$, représente la proportion des individus de la population pour lesquels la valeur du caractère est inférieure ou égale à x.

2. Caractère quantitatif continu

- Dans le cas continu, on suppose que le cumul s'effectue de façon linéaire entre le début et la fin de chaque classe.


G(x)
Diagramme cumulatif
N_i
N/2
N_{i-1}
x_i
M_e
x_{i+1}
x

2. Caractère quantitatif continu

Ainsi, la fonction de répartition des effectifs $G(x)$, est définie de $\mathbb{R}$ vers $[0, N]$, par :

$$
\begin{array}{lll}
G(x) = 0 & \text{si} & x \leq x_1 \\
G(x) = N_{i-1} + \dfrac{N_i - N_{i-1}}{x_{i+1} - x_i}(x - x_i) & \text{si} & x \in [x_i; x_{i+1}];\ 1 \leq i \leq k \\
G(x) = N & \text{si} & x \geq x_{k+1}
\end{array}
$$

avec

$$N_i = G(x_{i+1}),\ \ 1 \leq i \leq k$$

$$N_0 = G(x_1) = 0$$

2. Caractère quantitatif continu

La courbe cumulative (ou la courbe des effectifs cumulés) est la représentation graphique de la fonction $G(x)$.

2. Caractère quantitatif continu

Exemple2 : Taille des étudiants

Taille (en cm)	Effectif n_i	Eff. cum. N_i
$[151; 155[$	10	10
$[155; 159[$	12	22
$[159; 163[$	11	33
$[163; 167[$	7	40
$[167; 171[$	10	50
Total	50	

2. Caractère quantitatif continu

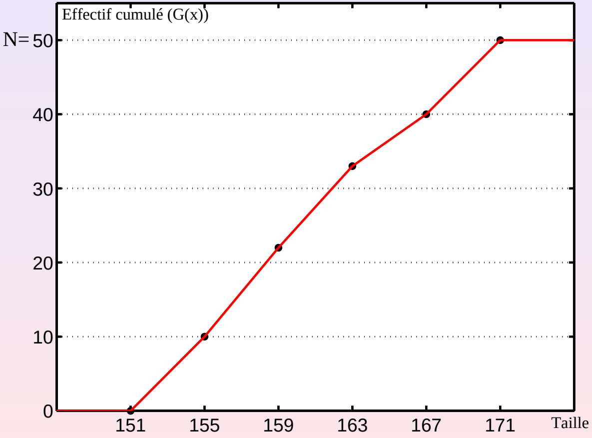


Figure: Courbe cumulative.

2. Caractère quantitatif continu

Remarque

- La courbe cumulative est construite en joignant par des segments de droites les points de coordonnées (x_{i+1}, N_i), $0 \leq i \leq k$.

2. Caractère quantitatif continu

Remarque :

La courbe cumulative est aussi la représentation graphique de la fonction de répartition $F(x)$, définie de $\mathbb{R}$ vers $[0, 1]$ par :

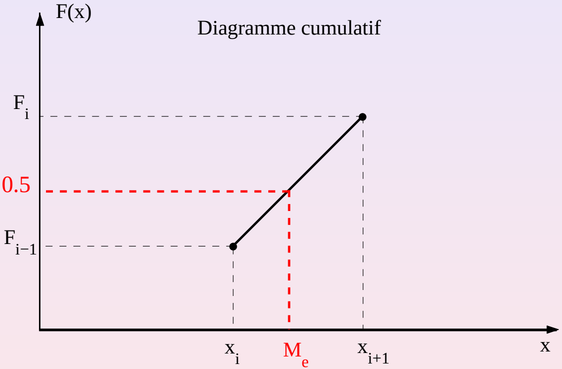
$$
\begin{array}{lll}
F(x) = 0 & \text{si} & x \leq x_1 \\
F(x) = F_{i-1} + \dfrac{F_i - F_{i-1}}{x_{i+1} - x_i}(x - x_i) & \text{si} & x \in [x_i; x_{i+1}];\ 1 \leq i \leq k \\
F(x) = 1 & \text{si} & x \geq x_{k+1}
\end{array}
$$

avec

$$F_0 = F(x_1) = 0$$

$$F_i = F(x_{i+1}),\ \ 1 \leq i \leq k$$

.


F(x)
Diagramme cumulatif
F_i
0.5
F_{i-1}
x_i
M_e
x_{i+1}
x

2. Caractère quantitatif continu

Remarque :

La courbe cumulative peut être construite en joignant par des segments de droites les points de coordonnées (x_{i+1}, F_i), $0 \leq i \leq k$.

2. Caractère quantitatif continu

Exemple2 : Taille des étudiants

Taille (en cm)	Effectif n_i	Eff. cum. N_i	Fréq. cum. F_i
$[151; 155[$	10	10	0.2
$[155; 159[$	12	22	0.44
$[159; 163[$	11	33	0.66
$[163; 167[$	7	40	0.8
$[167; 171[$	10	50	1
Total	50		

2. Caractère quantitatif continu

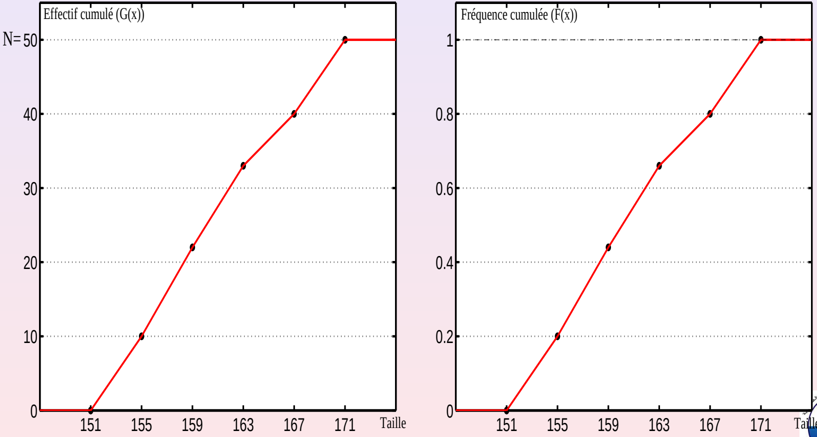


Figure: Courbe cumulative

2. Caractère quantitatif continu

Remarque :

- $G(x)$ est une fonction continue, croissante et passe progressivement de 0 à N
- $F(x)$ est une fonction continue, croissante et passe progressivement de 0 à 1
- On a :

$$F(x) = \frac{G(x)}{N}$$

III- Réduction des données statistiques

- Les tableaux statistiques et les représentations graphiques donnent une vue globale et détaillée de la distribution d'un caractère dans une population.
- Le but de la statistique descriptive est aussi de réduire et résumer les données d'une série par un ou plusieurs paramètres représentatifs et indiquer comment les valeurs de la série se répartissent autour de ces paramètres.

Il existe quatre types de paramètres :

1. Paramètres de tendance centrale (ou de position)
2. Paramètres de dispersion
3. Paramètres de forme
4. Paramètres de concentration

Les tableaux et graphiques contiennent la totalité des données. ils sont parfois durs à interpréter.

On cherchera alors à résumer les données statistiques par quelques paramètres numériques.

III.1- Paramètres de position

Appelés aussi paramètres de tendance centrale. Ils permettent de savoir autour de quelles valeurs se situent les données d'une série statistique.

1- Le mode

Definition

Le mode d'une série statistique, noté M_o, est la valeur du caractère qui admet le plus grand effectif.
C'est la valeur du caractère la plus fréquente.

Remarque :

- Pour un caractère continu, on définit d'abord la classe modale qui correspond à l'effectif corrigé le plus élevé.

Détermination pratique :

a. Cas d'un caractère discret :

Le mode correspond tout simplement à l'effectif le plus élevé. Il correspond au maximum du diagramme en bâtons.

Exemple1 :

Nombre d'enfants	Effectif n_i
0	4
1	10
2	12
3	8
4	5
5	4
6	4
7	2
8	1
Total	50

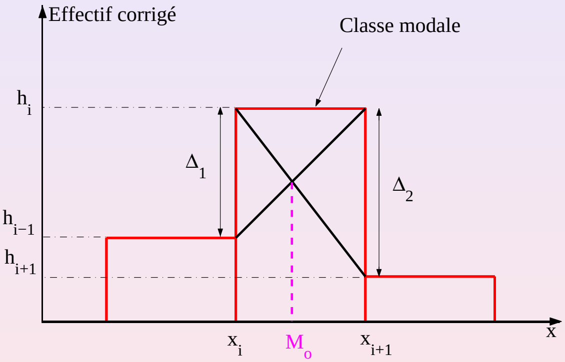
$$M_o = 2$$

b. Cas d'un caractère continu :

On définit d'abord la classe modale qui correspond à l'effectif corrigé le plus élevé.
La classe modale correspond donc à la classe dont le rectangle sur l'histogramme a la hauteur la plus élevée.

La classe modale $[x_i, x_{i+1}[$ étant déterminée, On détermine la valeur du mode M_o en tenant compte des effectifs corrigés des deux classes adjacentes à la classe modale par la méthode suivante :

Histogramme



On a : $$\frac{M_o - x_i}{\Delta_1} = \frac{x_{i+1} - M_o}{\Delta_2}$$

On a donc :

$$M_o = x_i + (x_{i+1} - x_i)\frac{\Delta_1}{\Delta_1 + \Delta_2}$$

avec

$$\Delta_1 = h_i - h_{i-1}$$

$$\Delta_2 = h_i - h_{i+1}$$

Exemple2 : Taille des étudiants

Taille (en cm)	Effectif n_i
$[151; 155[$	10
$[155; 159[$	12
$[159; 163[$	11
$[163; 167[$	7
$[167; 171[$	10
Total	50

Pour cet exemple, toutes les classes ont la même amplitude.
Donc on a : $h_i = n_i$

La classe modale est : $[155, 159[$.
On a :

$$\Delta_1 = 12 - 10 = 2$$

$$\Delta_2 = 12 - 11 = 1$$

d'où $\quad M_o = 155 + 4\dfrac{2}{2+1} = 157.66$ cm

Exemple2 : Taille des étudiants (les classes n'ont pas toutes la même amplitude)

Taille (en cm)	Effectif n_i	Amplitude a_i (en cm)	Eff. corrigé h_i
$[151; 155[$	10	4	10
$[155; 159[$	12	4	12
$[159; 167[$	18	8	9
$[167; 171[$	10	4	10
Total	50		

La classe modale est $[155; 159[$ (c'est la classe qui admet le plus grand effectif corrigé). On a :

$$\Delta_1 = 12 - 10 = 2$$

$$\Delta_2 = 12 - 9 = 3$$

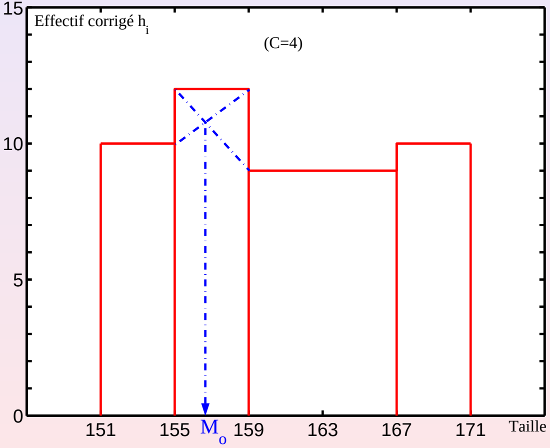
Donc

$$M_o = 155 + 4\frac{2}{2+3}$$

Le mode de cette distribution est donc :

$$M_o = 156.6\ cm$$

Détermination graphique :



2- La médiane

Definition

La médiane d'une série statistique ordonnée est la plus petite valeur x telle que $G(x) \geq \frac{N}{2}$ (ou $F(x) \geq 0,5$). On la note M_e.

G (resp. F) désigne la fonction des effectifs cumulés (resp. fréquences cumulées).

Interprétation

- La médiane M_e est la valeur centrale de la série statistique ordonnée.
  Elle partage la population en deux sous populations de même effectif.

  plus de 50% de la population ont des valeurs $\leq$ à M_e et
  plus de 50% de la population ont des valeurs $\geq$ à M_e.

Détermination pratique :

a. Cas d'un caractère discret :

Soient $x_{(1)} \leq x_{(2)} \leq \cdots \leq x_{(N)}$ les valeurs, ordonnées par ordre croissant, prises par les N individus d'une population. Ces valeurs ne sont pas nécessairement toutes distinctes.

- Si N est impair, la médiane est la valeur centrale de rang $\frac{N+1}{2}$:

$$M_e = x_{\left(\frac{N+1}{2}\right)}$$

- Si N est pair, la médiane est entre les deux valeurs centrales de rangs respectifs $\frac{N}{2}$ et $\frac{N}{2} + 1$. On prend :

$$M_e = \frac{1}{2}\left(x_{\left(\frac{N}{2}\right)} + x_{\left(\frac{N}{2}+1\right)}\right)$$

Exemple :

Soit un échantillon de 10 personnes dont les poids (en Kg) sont :

$$45 - 68 - 74 - 68 - 62 - 56 - 45 - 52 - 63 - 55$$

Exemple :

Soit un échantillon de 10 personnes dont les poids (en Kg) sont :

$$45 - 68 - 74 - 68 - 62 - 56 - 45 - 52 - 63 - 55$$

La série ordonnée par ordre croissant :

$$\underbrace{45 - 45 - 52 - 55 - 56}_{5} \qquad \uparrow \qquad \underbrace{62 - 63 - 68 - 68 - 74}_{5}$$

médiane

La médiane est donc :

$$M_e = \frac{56 + 62}{2} = 59 \text{ Kg}$$

Généralement, les données statistiques sont résumées par un tableau statistique. Soient $x_1, x_2, \ldots, x_k$ les modalités, ordonnées par ordre croissant, du caractère et Soient $N_1, N_2, \ldots, N_k$ les effectifs cumulés correspondants (on a $N_i = G(x_i)$).

- Si $N_i < \frac{N}{2} < N_{i+1}$, alors la médiane est la modalité d'effectif cumulé N_{i+1} :

$$M_e = x_{i+1}$$

on a $G(M_e) = G(x_{i+1}) = N_{i+1} > \frac{N}{2}$

- Si $\frac{N}{2} = N_i$, alors la médiane est le milieu entre les deux modalités d'effectif cumulé respectifs N_i et N_{i+1} :

$$M_e = \frac{1}{2}(x_i + x_{i+1})$$

on a $G(M_e) = G(x_i) = N_i = \frac{N}{2}$

Exemple2 :

Nombre d'enfants	Effectif n_i	Eff. cum. N_i
0	4	4
1	10	14
2	12	26
3	8	34
4	5	39
5	4	43
6	4	47
7	2	49
8	1	50

On a $14 < \frac{N}{2} = 25 < 26$, donc :

$$M_e = G^{-1}(26) = 2$$

Remarque :

On a le même raisonement avec les fréquences cumulées :

- Si $0.5 = F_i$, alors la médiane est le milieu entre les deux modalités de fréquence cumulée respectives F_i et F_{i+1} :

$$M_e = \frac{1}{2}(x_i + x_{i+1})$$

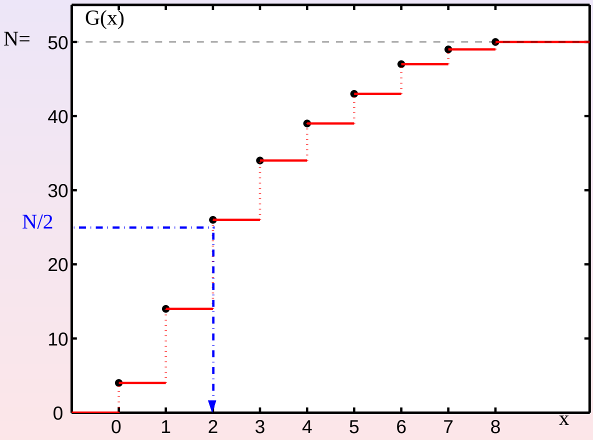
on a $F(M_e) = F_i = 0.5$

- Si $F_i < 0.5 < F_{i+1}$, alors la médiane est la modalité de fréquence cumulée F_{i+1} :

$$M_e = x_{i+1}$$

on a $F(M_e) = F_{i+1} > 0.5$

Détermination graphique :



b. Cas d'un caractère continu :

Pour des données groupées en classes, on situe d'abord la médiane à l'intérieur d'une classe $[x_i, x_{i+1}[$, appelée classe médiane.
C'est la première classe où l'effectif cumulé dépasse $\frac{N}{2}$ (ou la fréquence cumulée dépasse 0.5) :

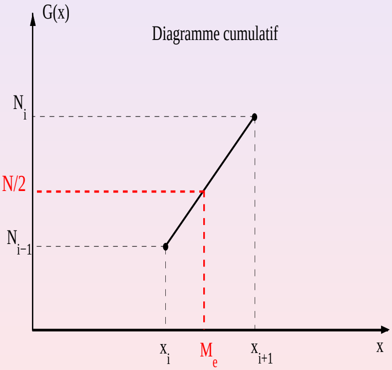
$$M_e \in [x_i, x_{i+1}[\Leftrightarrow N_{i-1} \leq \frac{N}{2} < N_i \Leftrightarrow F_{i-1} \leq 0.5 < F_i$$

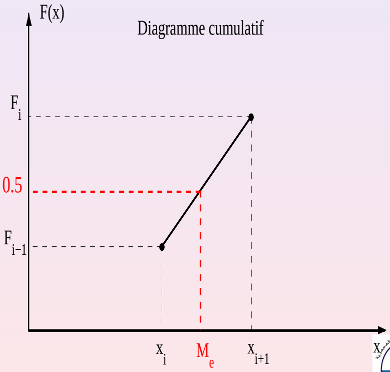
On a :

$$G(M_e) = \frac{N}{2}, \quad G(x_i) = N_{i-1}, \quad G(x_{i+1}) = N_i$$

$$F(M_e) = 0.5, \quad F(x_i) = F_{i-1}, \quad F(x_{i+1}) = F_i$$

Ensuite, la médiane est calculée par interpolation linéaire :





$$\frac{M_e - x_i}{x_{i+1} - x_i} = \frac{N/2 - N_{i-1}}{N_i - N_{i-1}} = \frac{0.5 - F_{i-1}}{F_i - F_{i-1}}$$

$$\Downarrow$$

$$M_e = x_i + (x_{i+1} - x_i)\frac{N/2 - N_{i-1}}{N_i - N_{i-1}}$$

ou

$$M_e = x_i + (x_{i+1} - x_i)\frac{0.5 - F_{i-1}}{F_i - F_{i-1}}$$

Exemple2 : Taille des étudiants

Taille (en cm)	Effectif n_i	Eff. cum. N_i	Fréquence f_i	Fréq. cum. F_i
[151; 155[	10	10	0.2	0.2
[155; 159[	12	22	0.24	0.44
[159; 163[	11	33	0.22	0.66
[163; 167[	7	40	0.14	0.8
[167; 171[	10	50	0.2	1
Total	50		1	

On a $\frac{N}{2} = 25$. Donc :

$$159 \to 22$$

$$M_e \to 25$$

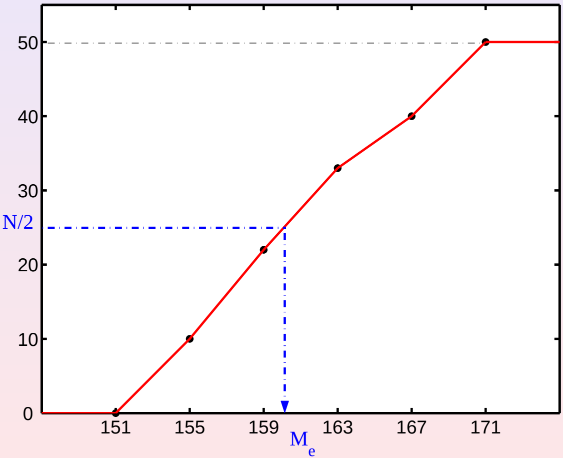
$$163 \to 33$$

d'où

$$M_e = 159 + (163 - 159)\frac{25 - 22}{33 - 22} = 160.09\ cm$$

50 % des étudiants ont une taille $\leq 160.09\ cm$ et 50 % des étudiants ont une taille $\geq 160.09\ cm$.

Détermination graphique :



3- La moyenne

Definition

La moyenne arithmétique d'un caractère statistique X, qu'on appelle tout simplement moyenne, est égale à la somme de toutes valeurs observées divisée par le nombre d'observations.

On la note $\overline{X}$.

Détermination pratique :

a. Cas d'un caractère discret :

Soit $(x_i, n_i)_{1 \leq i \leq k}$ une série statistique discrète. On a :

$$\overline{X} = \frac{1}{N} \sum_{i=1}^{k} n_i x_i$$

b. Cas d'un caractère continu :

Soit $([x_i, x_{i+1}[, n_i)_{1 \leq i \leq k}$ une série statistique continue. On a :

$$\overline{X} = \frac{1}{N} \sum_{i=1}^{k} n_i c_i$$

avec $c_i = \dfrac{x_i + x_{i+1}}{2} =$ centre de la classe $[x_i, x_{i+1}[$.

Cela suppose que toutes les observations à l'intérieur d'une classe sont groupées en son centre.

Exemple

$[x_i, x_{i+1}[$	n_i	c_i	$n_i \times c_i$
$[200, 300[$	8	250	2000
$[300, 400[$	26	350	9100
$[400, 500[$	12	450	5400
$[500, 600[$	10	550	5500
$[600, 700[$	15	650	9750
$[700, 800[$	5	750	3750
$[800, 900[$	3	850	2550
$[900, 1000[$	2	950	1900
Total	81		39950

Donc la moyenne est :

$$\overline{X} = \frac{39950}{81} = 493.21$$

Propriété :

- Si on pose $Y = aX + b$ (a et b deux constantes) alors $\overline{Y} = a\overline{X} + b$
- La somme des différences à la moyenne est nulle : $\sum_{i=1}^{k} n_i(x_i - \overline{X}) = 0$
- L'expression $\sum_{i=1}^{k} n_i(x_i - a)^2$ est minimale pour $a = \overline{X}$

Exemple : Distribution des salaires.

Effectuons le changement de variables $Y = \dfrac{X - 6500}{1000}$ pour simplifier les calculs :

Salaire en dh	n_i	c_i	$y_i = \dfrac{c_i - 6500}{1000}$	$n_i y_i$
$[3000, 5000[$	26	4000	-2.5	-65
$[5000, 6000[$	33	5500	-1	-33
$[6000, 7000[$	64	6500	0	0
$[7000, 8000[$	7	7500	1	7
$[8000, 10000[$	10	9000	2.5	25
Total	140			-66

$$\Downarrow$$

$$\overline{Y} = \frac{1}{N}\sum_{i=1}^{5} n_i y_i = -\frac{66}{140}$$

Et on a $\quad X = 1000\,Y + 6500$

D'où

$$\overline{X} = 1000\,\overline{Y} + 6500 = 6028.57\ dh$$

Propriétés :

- Si une population P de taille N est composée de m sous-populations $P_1, P_2, \ldots, P_m$, de tailles respectives $N_1, N_2, \ldots, N_m$ et de moyennes respectives $\overline{X}_1, \overline{X}_2, \ldots, \overline{X}_m$. Alors la moyenne $\overline{X}$ de la population P est donnée par :

$$\overline{X} = \frac{1}{N} \sum_{i=1}^{m} N_i \overline{X}_i$$

- Si X et Y sont deux caractès définis sur une même population P, alors :

$$\overline{X + Y} = \overline{X} + \overline{Y}$$

II.2- Paramètres de dispersion

Les paramètres de position ne donnent pas une information complète sur une variable statistique.

Deux variables qui ont les mêmes paramètres de position peuvent se présenter avec des dispersions très différentes.

Notes 1 : 7, 8, 11, 12, 13, 13, 13
Notes 2 : 4, 7, 9, 12, 13, 13, 19

1- L'étendue

Definition

L'étendue d'un caractère statistique X, noté $e(X)$, est la différence entre la plus grande valeur et la plus petite valeur observées :

$$e(X) = X_{max} - X_{min}$$

Ce paramètre présente un intérêt très limité du fait qu'il dépend uniquement des valeurs extrêmes, qui peuvent être des valeurs aberrantes.

2- Ecart absolu moyen

Definition

L'écart absolu moyen d'un caractère statistique X, noté $E_m(X)$, est la moyenne arithmétique des valeurs absolues des écarts à la moyenne.

$$E_m(X) = \frac{1}{N}\sum_{i=1}^{k} n_i|x_i - \overline{X}|$$

les x_i représentent les modalités dans le cas discret ou les centres des classes dans le cas continu.

Ce paramètre est peu maniable à cause des valeurs absolues.

3- Variance, Ecart-type

Definition

La variance d'un caractère statistique X, notée $V(X)$, est la moyenne arithmétique des carrés des écarts à la moyenne.

$$V(X) = \frac{1}{N}\sum_{i=1}^{k} n_i(x_i - \overline{X})^2$$

Definition

L'écart-type, noté $\sigma(X)$, est la racine carrée de la variance:

$$\sigma(X) = \sqrt{V(X)}$$

Formule simplifiée de la variance :

Theorem

On a :

$$V(X) = \overline{X^2} - \overline{X}^2$$

La variance est égale à la différence entre la moyenne des carrées et le carré de la moyenne.

$$\overline{X^2} = \frac{1}{N}\sum_{i=1}^{k} n_i x_i^2$$

Remarques :

- l'écart-type est le meilleur indicateur de la dispersion d'une série statistique par rapport à sa moyenne. Elle tient compte de toutes les valeurs de la série statistique.
- plus l'écart-type est faible, plus les valeurs du caractère sont concentrées autour de la moyenne.
- plus l'écart-type est élevé, plus les valeurs du caractère sont dispersées autour de la moyenne.
- La variance (ou l'écart-type) est nulle si et seulement si toutes les valeurs sont identiques et égales à la moyenne.

Propriétés :

Si on pose $Y = aX + b$ (a et b deux constantes) alors :

- $V(Y) = a^2 V(X)$
- $\sigma(Y) = |a|\sigma(X)$

Exemple : Distribution des salaires.

Salaire	c_i	n_i	$y_i = \frac{c_i - 6500}{1000}$	$n_i y_i$	$n_i y_i^2$
$[3000, 5000[$	4000	26	-2.5	-65	162.5
$[5000, 6000[$	5500	33	-1	-33	33
$[6000, 7000[$	6500	64	0	0	0
$[7000, 8000[$	7500	7	1	7	7
$[8000, 10000[$	9000	10	2.5	25	62.5
Total	140			-66	265

Effectuons le changement de variables $Y = \frac{X - 6500}{1000}$ pour simplifier les calculs :

$$\overline{Y} = \frac{1}{N}\sum_{i=1}^{5} n_i y_i = -\frac{66}{140}$$

$$\Rightarrow \quad \overline{X} = 1000\overline{Y} + 6500 = 6028.57$$

$$V(Y) = \frac{1}{N}\sum_{i=1}^{5} n_i y_i^2 - \overline{Y}^2 = \frac{265}{140} - (\frac{66}{140})^2 = 1.670612$$

$$\Rightarrow \quad V(X) = 1000^2 V(Y) = 1670612$$

$$\Rightarrow \quad \sigma(X) = \sqrt{V(X)} = 1292,52$$

4- Quartiles Ecart interquartile

Définition

Le premier quartile d'une série statistique ordonnée est la plus petite valeur x telle que $G(x) \geq \frac{N}{4}$ (ou $F(x) \geq 0.25$). On le note Q_1.

Le troisième quartile d'une série statistique ordonnée est la plus petite valeur x telle que $G(x) \geq \frac{3N}{4}$ (ou $F(x) \geq 0.75$). On le note Q_3.

Interprétation

- Plus de 25% de la population ont des valeurs $\leq$ à Q_1 et plus de 75% de la population ont des valeurs $\geq$ à Q_1.
- Plus de 75% de la population ont des valeurs $\leq$ à Q_3 et plus de 25% de la population ont des valeurs $\geq$ à Q_3.

Remarque

La médiane est le deuxième quartile :

$$M_e = Q_2$$

Les quartiles sont des paramètres de position.
C'est l'écart inter-quartile qui est un paramétre de dispersion.

Détermination pratique d'un quartile Q_p :

a. Cas d'un caractère discret :

Soient $x_{(1)} \leq x_{(2)} \leq \cdots \leq x_{(N)}$ les valeurs, ordonnées par ordre croissant, prises par les N individus d'une population. Ces valeurs ne sont pas nécessairement toutes distinctes.

- Si $p \times \frac{N}{4}$ est un nombre entier, alors

$$Q_p = \frac{1}{2}\left(x_{(p\frac{N}{4})} + x_{(p\frac{N}{4}+1)}\right)$$

- Si $p \times \frac{N}{4}$ n'est pas un nombre entier, alors

$$Q_p = x_{([p\frac{N}{4}])}$$

où $[p\frac{N}{4}]$ représente le plus petit nombre entier supérieur à $p\frac{N}{4}$.

Exemple:
Soit la série statistique ordonnée :
12, 13, 15, 16, 18, 19, 22, 24, 25, 27. On a :

Exemple:
Soit la série statistique ordonnée :
12, 13, 15, 16, 18, 19, 22, 24, 25, 27. On a :

$$Q_1 = x_{([2.5])} = x_{(3)} = 15$$

$$Q_2 = \frac{1}{2}(x_{(5)} + x_{(6)}) = \frac{(18+19)}{2} = 18.5$$

$$Q_3 = x_{([7.5])} = x_{(8)} = 24$$

Soient $x_1, x_2, \ldots, x_k$ les modalités, ordonnées par ordre croissant, du caractère et Soient $N_1, N_2, \ldots, N_k$ les effectifs cumulés correspondants (on a $N_i = G(x_i)$).

- Si $p \times \frac{N}{4} = N_i$, alors le quartile Q_p est le milieu entre les deux modalités d'effectif cumulé respectifs N_i et N_{i+1} :

$$Q_p = \frac{1}{2}(x_i + x_{i+1})$$

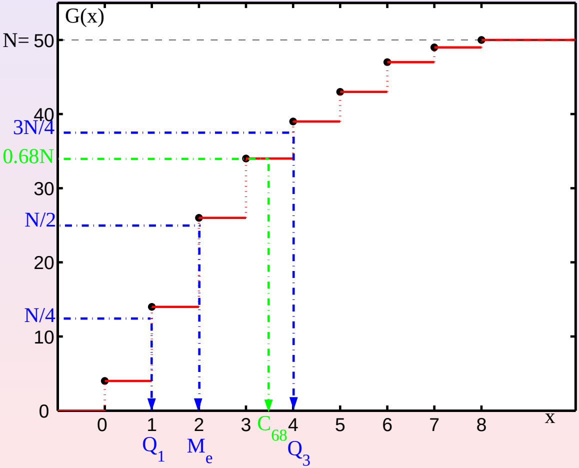
- Si $N_i < p \times \frac{N}{4} < N_{i+1}$, alors le quartile Q_p est la modalité d'effectif cumulé N_{i+1} :

$$Q_p = x_{i+1}$$

Exemple2 :

Nombre d'enfants	Effectif n_i	Eff. cum. N_i
0	4	4
1	10	14
2	12	26
3	8	34
4	5	39
5	4	43
6	4	47
7	2	49
8	1	50
Total	50	

Détermination graphique :



Détermination pratique du quartile Q_p : Cas continu

- Méthode d'interpolation : d'après le tableau statistique ou la courbe cumulative, on détermine d'abord la classe $[x_i, x_{i+1}[$ telle que :

$$N_{i-1} \leq p \times \frac{N}{4} < N_i$$

Puis, par interpolation linéaire, on calcule Q_p :

$$Q_p = x_i + (x_{i+1} - x_i)\frac{pN/4 - N_{i-1}}{N_i - N_{i-1}}$$

Exemple:

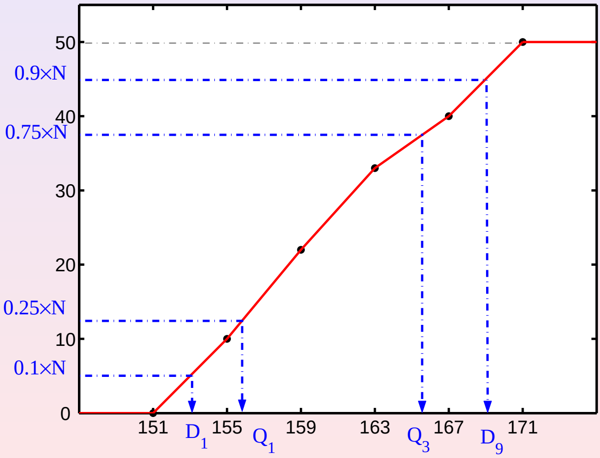
Taille (en cm)	Effectif n_i	Eff. cum. N_i
$[151; 155[$	10	10
$[155; 159[$	12	22
$[159; 163[$	11	33
$[163; 167[$	7	40
$[167; 171[$	10	50
Total	50	

$$Q_1 = 155 + 4\frac{12.5 - 10}{22 - 10} = 155.83\ cm$$

$$Q_3 = 163 + 4\frac{37.5 - 33}{40 - 33} = 165.57\ cm$$

- Méthode graphique : On trace la courbe cumulative, et on détermine Q_p comme l'abscisse du point d'ordonnée $p \times \frac{N}{4}$

Détermination graphique :



Remarques :

- On a le même raisonement avec les fréquences cumulées.
- Après les quartiles, on peut définir de la même façon les déciles (voire les centiles) d'une série statistique.

L'écart interquartiles

Définition

L'écart interquartile est la différence entre le dernier et le premier quartile :

$$\Delta Q = Q_3 - Q_1$$

L'intervalle interquantile est l'intervalle :

$$[Q_1, Q_3]$$

Interprétation :

- L'écart interquartile mesure la dispersion sans tenir compte des valeurs extrêmes.
- L'intervalle interquartile contient au moins 50% des valeurs de la série.
- Plus l'intervalle interquantile est large, plus la série est dispersée.

5- Paramètres de dispersion relative

Remarque

- La dispersion mesurée par les paramètres présentés précédemment est qualifiée d'absolue : ils s'exprime dans l'unité de mesure du caractère.

- Pour comparer la dispersion de deux séries statistiques ayant des unités différentes (ou même des ordres de grandeur différents), il faut considérer des paramètres de dispersion relative.

Définition

Coefficient de variation :

$$C_v = \frac{\sigma(X)}{\overline{X}}$$

Ecart moyen relatif :

$$\frac{E_m(X)}{\overline{X}}$$

Coefficient interquartile relatif :

$$\frac{Q_3 - Q_1}{M_e}$$

Fin chapitre 1

Chapitre 2

Statistique descriptive à deux dimensions

Sommaire

Introduction

Très souvent, on peut envisager pour une même population l'étude de deux ou plusieurs caractères considérés simultanément.

Exemple :

- Notes de Mathématique et notes de Physique d'un groupe d'étudiants.
- Nombre d'années d'ancienneté et revenus mensuels des employés d'une entreprise.

Le but essentiel de ce chapitre est de :

- Étendre les notions de la statistique descriptive à une dimension au cas d'un couple de caractères (deux dimensions).
- Caractériser les relations de dépendance qui peuvent exister entre les deux caractères.

Considérons une population P de N individus pour chacun desquels on détermine deux valeurs relatives à deux caractères X et Y.

La suite $(x_{(1)}, y_{(1)}), (x_{(2)}, y_{(2)}), \ldots, (x_{(N)}, y_{(N)})$, des N couples de valeurs prises par les deux caractères X et Y, s'appelle série statistique brute associée au couple de caractères (X, Y).

Les modalités de X sont notées :

$$x_1, x_2, \ldots, x_k$$

Les modalités de Y sont notées :

$$y_1, y_2, \ldots, y_p$$

Toute valeur (x_i, y_j) s'appelle modalité du couple (X, Y).

On appelle effectif de la modalité (x_i, y_j), (noté n_{ij}), le nombre d'individus présentant cette modalité. On a :

$$\sum_{i=1}^{k} \sum_{j=1}^{p} n_{ij} = N$$

La famille $((x_i, y_j), n_{ij})$ est appelée série statistique double (à deux dimensions) associée au couple de caractères (X, Y).

Remarques

- Les effectifs cumulés ne sont pas définis, car il n'existe pas d'ordre sur $\mathbf{R}^2$.
- On peut procéder à des regroupement en classes pour l'un des caractères ou pour les deux. On peut ainsi, définir des séries doubles semi-continues (un seul caracère est continu) ou continues (les deux caractères sont continus).

I- Présentation des données

Les données statistiques relatives à un couple de caractès X et Y, sont généralement, présentées sous forme de distributions d'effectifs ou de fréquences dans un tableau statistique à deux dimensions, appelé tableau de contingence.

A partir de ce tableau, on peut déterminer :

- la distribution conjointe
- les distributions marginales
- les distributions conditionnelles

Tableau des effectifs

$X \backslash Y$	y_1	y_2	$\dots$	y_j	$\dots$	y_p	Total
x_1	n_{11}	n_{12}	$\dots$	n_{1j}	$\dots$	n_{1p}	$n_{1.}$
x_2	n_{21}	n_{22}	$\dots$	n_{2j}	$\dots$	n_{2p}	$n_{2.}$
$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$	$\vdots$
x_i	n_{i1}	n_{i2}	$\dots$	n_{ij}	$\dots$	n_{ip}	$n_{i.}$
$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$	$\vdots$
x_k	n_{k1}	n_{k2}	$\dots$	n_{kj}	$\dots$	n_{kp}	$n_{k.}$
Total	$n_{.1}$	$n_{.2}$	$\dots$	$n_{.j}$	$\dots$	$n_{.p}$	N

- n_{ij} : effectif de la modalité (x_i, y_j) du couple (X, Y)

- $n_{i.}$: effectif de la modalité x_i de X ($\forall$ la valeur de Y).

- $n_{.j}$: effectif de la modalité y_j de Y ($\forall$ la valeur de X).

Les n_{ij} sont appelés effectifs conjoints du couple (X, Y).

Les $n_{i.}$ et $n_{.j}$ sont appelés effectifs marginaux (respectivement de X et de Y).

On a les relations :

$$\sum_{j=1}^{p} n_{ij} = n_{i.}$$

$$\sum_{i=1}^{k} n_{ij} = n_{.j}$$

$$\sum_{i=1}^{k} n_{i.} = \sum_{j=1}^{p} n_{.j} = \sum_{i=1}^{k} \sum_{j=1}^{p} n_{ij} = N$$

Tableau des fréquences

Le tableau de fréquences conjointes et marginales s'obtient en divisant tous les effectifs par la taille de la poulation :

$X \backslash Y$	y_1	y_2	$\dots$	y_j	$\dots$	y_p	Total
x_1	f_{11}	f_{12}	$\dots$	f_{1j}	$\dots$	f_{1p}	$f_{1.}$
x_2	f_{21}	f_{22}	$\dots$	f_{2j}	$\dots$	f_{2p}	$f_{2.}$
$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$	$\vdots$
x_i	f_{i1}	f_{i2}	$\dots$	f_{ij}	$\dots$	f_{ip}	$f_{i.}$
$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$	$\vdots$
x_k	f_{k1}	f_{k2}	$\dots$	f_{kj}	$\dots$	f_{kp}	$f_{k.}$
Total	$f_{.1}$	$f_{.2}$	$\dots$	$f_{.j}$	$\dots$	$f_{.p}$	1

$$f_{ij} = \frac{n_{ij}}{N}$$

$$f_{i.} = \frac{n_{i.}}{N} = \sum_{j=1}^{p} f_{ij}$$

$$f_{.j} = \frac{n_{.j}}{N} = \sum_{i=1}^{k} f_{ij}$$

$$\sum_{i=1}^{k} f_{i.} = \sum_{j=1}^{p} f_{.j} = \sum_{i=1}^{k} \sum_{j=1}^{p} f_{ij} = 1$$

- f_{ij} représente la proportion des individus ayant la modalité (x_i, y_j) du couple (X, Y).

- $f_{i.}$ représente la proportion des individus ayant la modalité x_i de X ($\forall$ la valeur de Y).

- $f_{.j}$ représente la proportion des individus ayant la modalité y_j de Y ($\forall$ la valeur de X).

Les f_{ij} sont appelés fréquences conjointes du couple (X, Y).

Les $f_{i.}$ et $f_{.j}$ sont appelés fréquences marginales, respectivement de X et de Y.

Distribution conjointe

Définition

On appelle distribution conjointe du couple (X, Y) la distribution à deux dimensions des individus de la population en fonction des modalités (x_i, y_j) du couple (X, Y).
Elle correspond aux effectifs n_{ij} (ou aux fréquence f_{ij}) du tableau de contingence.

Distributions marginales

Définition

On appelle distribution marginale du caractère X la distribution à une dimension des individus de la population en fonction des modalités x_i de X, indépendamment de Y.
Elle correspond aux effectifs $n_{i.}$ (ou aux fréquences $f_{i.}$) de la dernière colonne du tableau de contingence.

De façon analogue, on définit la distribution marginale de Y.
Elle correspond aux effectifs $n_{.j}$ (ou aux fréquences $f_{.j}$) de la dernière ligne du tableau de contingence.

Distributions conditionnelles

Définition

On appelle distribution conditionnelle de X sous la condition $Y = y_j$, la distribution à une dimension des individus de la population en fonction des modalités x_i de X, sachant que la modalité de Y est y_j.
Elle correspond aux effectifs $n_{ij}, \quad 1 \leq i \leq k$ de la colonne j du tableau de contingence.

De façon analogue, on définit la distribution conditionnelle de Y sous la condition $X = x_i$.
Elle correspond aux effectifs $n_{ij}, \quad 1 \leq j \leq p$ de la ligne i du tableau de contingence.

Les distributions conditionnelles de X sachant Y sont présentées dans un tableau de profils colonnes :

$X \backslash Y$	y_1	y_2	$\dots$	y_j	$\dots$	y_p
x_1	n_{11}	n_{12}	$\dots$	n_{1j}	$\dots$	n_{1p}
x_2	n_{21}	n_{22}	$\dots$	n_{2j}	$\dots$	n_{2p}
$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$
x_i	n_{i1}	n_{i2}	$\dots$	n_{ij}	$\dots$	n_{ip}
$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$
x_k	n_{k1}	n_{k2}	$\dots$	n_{kj}	$\dots$	n_{kp}
Total	$n_{.1}$	$n_{.2}$	$\dots$	$n_{.j}$	$\dots$	$n_{.p}$

On a p distributions conditionnelles de X sachant Y

Les distributions conditionnelles de Y sachant X sont présentées dans un tableau de profils lignes :

$X \backslash Y$	y_1	y_2	$\dots$	y_j	$\dots$	y_p	Total
x_1	n_{11}	n_{12}	$\dots$	n_{1j}	$\dots$	n_{1p}	$n_{1.}$
x_2	n_{21}	n_{22}	$\dots$	n_{2j}	$\dots$	n_{2p}	$n_{2.}$
$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$	$\vdots$
x_i	n_{i1}	n_{i2}	$\dots$	n_{ij}	$\dots$	n_{ip}	$n_{i.}$
$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$	$\vdots$
x_k	n_{k1}	n_{k2}	$\dots$	n_{kj}	$\dots$	n_{kp}	$n_{k.}$

On a k distributions conditionnelles de Y sachant X

Remarques :

- Pour un couple de caractères (X, Y), on a :

  - une distribution conjointe,
  - deux distributions marginales,
  - $p + k$ distributions conditionnelles,

- les distributions marginales et les distributions conditionnelles sont des distributions à une seule dimension. Elles peuvent être étudiées conformément à ce que nous avons vu au chapitre 1.

Fréquences conditionnelles

Définition

La fréquence conditionnelle de la modalité x_i du caractère X sachant que $Y = y_j$, (notée f_{x_i/y_j}), est définie par :

$$f_{x_i/y_j} = \frac{n_{ij}}{n_{.j}} = \frac{f_{ij}}{f_{.j}}$$

Elle correspond à la proportion d'individus présentant la modalité x_i du caractère X dans le sous-groupe d'individus présentant la modalité y_j du caractère Y.

Fréquences conditionnelles

Définition

La fréquence conditionnelle du caractère Y sachant que $X = x_i$, (notée f_{y_j/x_i}), est définie par :

$$f_{y_j/x_i} = \frac{n_{ij}}{n_{i.}} = \frac{f_{ij}}{f_{i.}}$$

Elle correspond à la proportion d'individus présentant la modalité y_j du caractère Y dans le sous-groupe d'individus présentant la modalité x_i du caractère X.

Exemple :

Les distributions conjointe, marginales et conditionnelles (en effectifs) des notes de mathématique et de physique d'un groupe d'étudiants se présentent dans le tableau de contingence suivant :

Physique \ Math	$[0,7[$	$[7,10[$	$[10,20]$	Total
$[0,7[$	17	5	4	26
$[7,10[$	3	21	9	33
$[10,15[$	3	10	111	124
$[15,20]$	0	1	16	17
Total	23	37	140	200

- 8,5% (=17/200) des étudiants n'ont pas validé les deux modules (fréquence conjointe).
- 13% (26/200) des étudiants n'ont pas validé le module de Physique (fréquence marginale).
- 70% (=140/200) des étudiants ont validé le module de Math (fréquence marginale).
- 8% (=16/200) des étudiants ont une note $\geq$ à 15 en Physique et la moyenne en Math (fréquence conjointe).

- 73,9% (=17/23) des étudiants n'ayant pas validé le module de Math n'ont pas validé le module de Physique (fréquence conditionnelle).
- 65,4% (=17/26) des étudiants n'ayant pas validé le module de Physique n'ont pas validé le module de Math (fréquence conditionnelle).
- La proportion des étudiants n'ayant pas validé le module de Physique sachant qu'ils ont obtenu la moyenne en Math est 2,9% (=4/140) (fréquence conditionnelle).

Remarque :

- Les distributions des fréquences conjointes et marginales s'obtiennent en divisant les effectifs conjoints et marginaux par N
- Les distributions des fréquences conditionnelles de X sachant Y sont présentées dans un tableau de profils colonnes en divisant tous les effectifs d'une colonne j par l'effectif marginale $n_{.j}$ de cette colonne.
- Les distributions des fréquences conditionnelles de Y sachant X sont présentées dans un tableau de profils lignes en divisant tous les effectifs d'une ligne i par l'effectif marginale $n_{i.}$ de cette ligne.

Remarque :

Pour les distributions des fréquences, le tableau de contingence des effectifs s'éclate en un tableau des distributions des fréquences conjointes et marginales et deux tableaux des distributions des fréquences conditionnelles.

Tableau des fréquences conditionnelles des notes de Physique sachant les notes de Math :

Physique \ Math	$[0,7[$	$[7,10[$	$[10,20]$
$[0,7[$	$17/23 \simeq 73,9\%$	$13,5\%$	$2,9\%$
$[7,10[$	$3/23 \simeq 13\%$	$56,8\%$	$6,4\%$
$[10,15[$	$3/23 \simeq 13\%$	27%	$79,3\%$
$[15,20]$	$0/23 = 0\%$	$2,7\%$	$11,4\%$
Total	100%	100%	100%

Tableau des fréquences conditionnelles des notes de Math sachant les notes de Physique :

Physique \ Math	$[0,5[$	$[5,10[$	$[10,20]$	Total
$[0,7[$	$17/26 \simeq 65,4\%$	$5/26$	$4/26$	100%
$[7,10[$	$3/33 \simeq 9,1\%$	$21/33$	$9/33$	100%
$[10,15[$	$3/124 \simeq 2,4\%$	$10/124$	$111/124$	100%
$[15,20]$	$0/17 = 0\%$	$1/17$	$16/17$	100%

II- Paramètres statistiques

On distingue deux types de paramètres :

- Les paramètres qui concernent un seul caractère, ils caractérisent les distributions partielles (marginales et conditionnelles).
- Les paramètres qui décrivent les relations qui existent entre les deux caractères, ils caractérisent la distribution conjointe.

II.1- Paramètres des distributions partielles

Les distributions partielles (marginales et conditionnelles) sont des distributions à une seule dimension.

Elles peuvent être étudiées conformément à ce que nous avons vu au chapitre 1.

Moyennes et variances marginales :

$$\overline{X} = \frac{1}{N} \sum_{i=1}^{k} n_{i.} x_i$$

$$V(X) = \frac{1}{N} \sum_{i=1}^{k} n_{i.} (x_i - \overline{X})^2$$

$$\overline{Y} = \frac{1}{N} \sum_{j=1}^{p} n_{.j} y_j$$

$$V(Y) = \frac{1}{N} \sum_{j=1}^{p} n_{.j} (y_j - \overline{Y})^2$$

Moyennes et variances conditionnelles :

Moyennes et variances conditionnelles de X sachant que $Y = y_j$:

$$\overline{X}_j = \frac{1}{n_{.j}} \sum_{i=1}^{k} n_{ij} x_i, \quad 1 \leq j \leq p$$

$$V_j(X) = \frac{1}{n_{.j}} \sum_{i=1}^{k} n_{ij} (x_i - \overline{X}_j)^2, \quad 1 \leq j \leq p$$

Moyennes et variances conditionnelles de Y sachant que $X = x_i$:

$$\overline{Y}_i = \frac{1}{n_{i.}} \sum_{j=1}^{p} n_{ij} y_j, \quad 1 \leq i \leq k$$

$$V_i(Y) = \frac{1}{n_{i.}} \sum_{j=1}^{p} n_{ij} (y_j - \overline{Y}_i)^2, \quad 1 \leq i \leq k$$

Relations entre paramètres marginales et paramètres conditionnelles :

On a :

$$\overline{X} = \frac{1}{N}\sum_{i=1}^{k}\sum_{j=1}^{p} n_{ij}x_i = \frac{1}{N}\sum_{j=1}^{p} n_{.j}\overline{X}_j$$

$$\overline{Y} = \frac{1}{N}\sum_{j=1}^{p}\sum_{i=1}^{k} n_{ij}y_j = \frac{1}{N}\sum_{i=1}^{k} n_{i.}\overline{Y}_i$$

La moyenne marginale est égale à la moyenne des moyennes conditionnelles.

De même on montre que :

$$V(X) = \frac{1}{N}\sum_{j=1}^{p} n_{.j} V_j(X) + \frac{1}{N}\sum_{j=1}^{p} n_{.j}(\overline{X}_j - \overline{X})^2$$

$$V(Y) = \frac{1}{N}\sum_{i=1}^{k} n_{i.} V_i(Y) + \frac{1}{N}\sum_{i=1}^{k} n_{i.}(\overline{Y}_i - \overline{Y})^2$$

La variance marginale est égale à la moyenne des variances conditionnelles plus la variance des moyennes conditionnelles.

II.2- Covariance

Définition

La covariance d'un couple de variables (X, Y), notée $cov(X, Y)$, est définie par :

$$cov(X,Y) = \overline{(X - \overline{X})(Y - \overline{Y})} = \frac{1}{N}\sum_{i=1}^{k}\sum_{j=1}^{p} n_{ij}(x_i - \overline{X})(y_j - \overline{Y})$$

Propriétés :

- $cov(X,Y) = \overline{XY} - \overline{X}.\overline{Y} = \dfrac{1}{N}\sum_{i=1}^{k}\sum_{j=1}^{p} n_{ij}x_i y_j - \overline{X}.\overline{Y}$
- Si $X' = aX + b$ et $Y' = cY + d$ (a, b, c et d des constantes) alors : $cov(X', Y') = ac \times cov(X,Y)$
- $V(X) = cov(X,X)$
- $V(X+Y) = V(X) + V(Y) + 2cov(X,Y)$
- $|cov(X,Y)| \leq \sigma(X)\sigma(Y)$

Remarque :

- La covariance est positive ou négative selon que la relation entre les variables X et Y est croissante ou décroissante, c'est à dire selon que les deux variables varient dans le même sens ou en sens inverse.

Exemple

Physique \ Math	$[0,7[$	$[7,10[$	$[10,20]$	Total
$[0,7[$	17	5	4	26
$[7,10[$	3	21	9	33
$[10,20[$	3	11	127	141
Total	23	37	140	200

- La moyenne marginale de la note de Physique est :

$$\overline{X} = \frac{26 \times 3.5 + 33 \times 8.5 + 141 \times 15}{200} = \frac{2486.5}{200} = 12.43$$

- La moyenne marginale de la note de Math est :

$$\overline{Y} = \frac{23 \times 3.5 + 37 \times 8.5 + 140 \times 15}{200} = \frac{2495}{200} = 12.47$$

- La moyenne du produit XY est :

$$\overline{XY} = \frac{3.5 \times (17 \times 3.5 + 3 \times 8.5 + 3 \times 15) + 8.5 \times (5 \times 3.5 + 21 \ldots}{200}$$

$$\overline{XY} = \frac{33456}{200} = 167.28$$

- D'où

$$cov(X, Y) = 12.18$$

III- AJUSTEMENT

Soit $\left((x_i, y_j), n_{ij}\right)_{1\leq i\leq k, 1\leq j\leq p}$ une série double associée à un couple de caractères (X, Y).

On se pose les questions suivantes :

- Existe-t-il une relation de dépendance $Y = f(X)$ entre les deux variables statistiques X et Y ?
- Si on connaît la valeur de l'une des variables, peut-on estimer la valeur de l'autre variable ?
- Comment mesurer la précision de l'estimation ou de la modélisation $Y = f(X)$?

III.1- Nuage statistique

Pour répondre aux questions précédentes, on commence par représenter le nuage statistique formé par les points $M_{ij} = (x_i, y_j)$ dans un repère cartésien.

Remarque

- Si une modalité (x_i, y_j) a un effectif $n_{ij} > 1$, on écrit ce nombre à côté du point M_{ij}.
- Si les valeurs des variables sont groupées en classes, on considère les centres des classes.

Suivant l'allure du nuage statistique, on essaiera de déterminer l'équation d'une courbe C approchant au mieux le nuage statistique.

La solution de ce problème est particulièrement simple lorsque les point M_{ij} du nuage statistique semblent à peu près alignés.

On cherchera alors une droite (D) d'équation $Y = aX + b$ approchant au mieux le nuage de points.

Cette droite s'appelle droite d'ajustement linéaire de Y en fonction de X.

Remarque

Il est possible que l'allure du nuage statistique suggère un ajustement non linéaire : exponentiel, puissance, logarithmique, etc.

III.2- Ajustement linéaire - Méthode des moindres carrées

On suppose que les point M_{ij} du nuage statistique sont plus au moins alignés de sorte que le nuage suggère un ajustement linéaire.

Les points $M_{ij} = (x_i, y_j)$ du nuage statistique sont renumérotés $M_1, M_2, \dots, M_n$.
L'effectif de chaque point $M_k = (x_{(k)}, y_{(k)})$ est noté n_k.

Definition

L'ajustement linéaire ou affine, par la méthode des moindres carrées, consiste à ajuster la distribution de (X, Y) par une droite D d'équation $Y = aX + b$ telle que la somme

$$S = \sum_{k=1}^{n} n_k(ax_k + b - y_k)^2$$

soit minimale.

La droite D s'appelle droite de régression linéaire de Y en fonction de X par la méthode des moindres carrées.

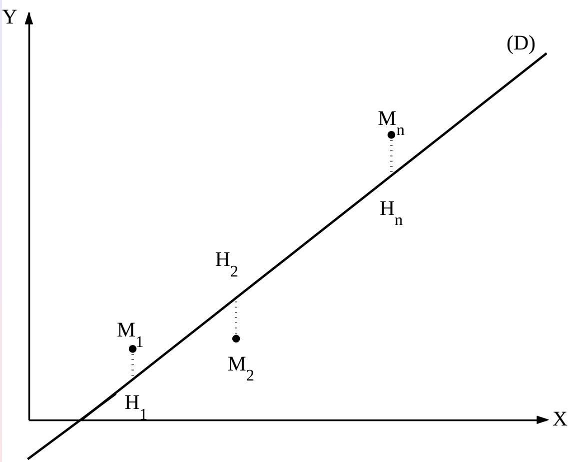


Figure: Méthode des moindres carrées.

Remarque

En échangeant les rôles de X et Y nous obtenons la droite de régression linéaire de X en fonction de Y.

Definition

La droite de régression linéaire de X en fonction de Y est la droite D' d'équation $X = a'Y + b'$ telle que la somme

$$S' = \sum_{k=1}^{n} n_k(a'y_k + b' - x_k)^2$$

soit minimale.

Théorème

- *La droite de régression D de Y en fonction de X, par la méthode des moindres carrés, a pour équation $Y = aX + b$ avec :*

$$a = \frac{Cov(X,Y)}{V(X)}, \quad b = \overline{Y} - a\overline{X}$$

- *La droite de régression D' de X en fonction de Y, par la méthode des moindres carrés, a pour équation $X = a'Y + b'$ avec :*

$$a' = \frac{Cov(X,Y)}{V(Y)}, \quad b' = \overline{X} - a'\overline{Y}$$

Remarque

Les droites de régression D et D' passent par le point moyen $G = (\overline{X}, \overline{Y})$.

Ainsi, pour tracer une droite de régression, il suffit de choisir un deuxième point appartenant à cette droite.

Exemple1

On considère un échantillon de 200 étudiants pour chacun desquels, on a relevé la note de mathématique X et la note de physique Y :

Physique (Y) / Math (X)	$[0,6[$	$[6,10[$	$[10,20]$	Total
$[0,6[$	18	9	0	27
$[6,10[$	3	21	9	33
$[10,12[$	2	8	77	87
$[12,16[$	0	3	34	37
$[16,20]$	0	0	16	16
Total	23	41	136	200

On commence par représenter le nuage des points $M_{ij}(x_i, y_j)$ dans un repère cartésien.

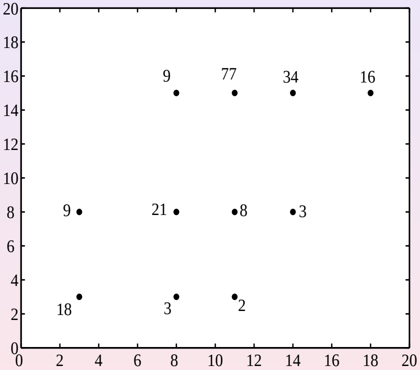


Figure: Nuage de points (x_i, y_j).

On va déterminer les droite d'ajustement (D) et (D') par la méthode des moindres carrées. Pour cela, on complète le tableau précédent en remplaçant chaque classe par son centre.
Les nombres entre parenthèse représentent les produits $n_{ij}x_iy_j$.

x_i \ y_j	3	8	15	$n_{i.}$	$n_{i.}x_i$	$n_{i.}x_i^2$
3	18 (162)	9 (216)	0 (0)	27 (378)	81	243
8	3 (72)	21 (1344)	9 (1080)	33 (2496)	264	2112
11	2 (66)	8 (704)	77 (12705)	87 (13475)	957	10527
14	0 (0)	3 (336)	34 (7140)	37 (7476)	518	7252
18	0 (0)	0 (0)	16 (4320)	16 (4320)	288	5184
$n_{.j}$	23 (300)	41 (2600)	136 (25245)	200 (28145)	2108	25318
$n_{.j}y_j$	69	328	2040	2437		

D'où les résultats :

$$\overline{X} = \frac{2108}{200} = 10.54$$

$$\overline{Y} = \frac{2437}{200} = 12.185$$

$$V(X) = \frac{25318}{200} - (10.54)^2 = 15.4984$$

$$V(Y) = \frac{33431}{200} - (12.185)^2 = 18.6808$$

$$Cov(X, Y) = \frac{28145}{200} - 10.54 \times 12.185 = 12.2951$$

Donc les valeurs de a et b sont données par :

$$a = \frac{12.2951}{15.4984} = 0.7933$$

$$b = 12.185 - 0.7841 \times 10.54 = 3.8235$$

D'où l'équation de la droite d'ajustement de Y en fonction de X :

$$(D) : Y = 0.79X + 3.82$$

De même, on obtient l'équation de la droite d'ajustement de X en fonction de Y :

$$(D') : X = 0.66Y + 2.52$$

On trace les droites (D) et (D') avec le nuage statistique.

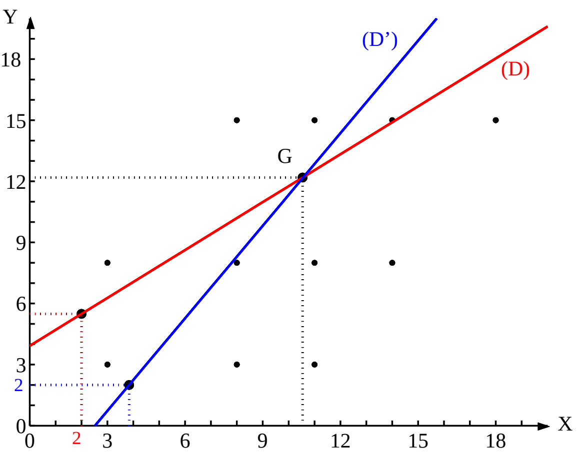


Figure: Droites de régression par la méthode des moindres carrés.

Exemple2

Cas particulier : Chaque modalité a un effectif $= 1$

Pour un couple de variables (X, Y) où chaque modalité (x_i, y_j) du couple a un effectif ≤ 1, les données peuvent être présentées sous forme d'un tableau de ce type :

X	x_1	x_2	$\ldots$	x_i	$\ldots$	x_N
Y	y_1	y_2	$\ldots$	y_i	$\ldots$	y_N

Exemple :
On considère un échantillon de 10 arbres pour chacun desquels, on a mesuré le diamètre X et la hauteur Y (en cm) :

X	15	21	17	22	20	13	19	15	18	23
Y	300	380	330	360	420	290	350	320	360	420

On commence par représenter le nuage des points (x_i, y_i) dans un repère cartésien.

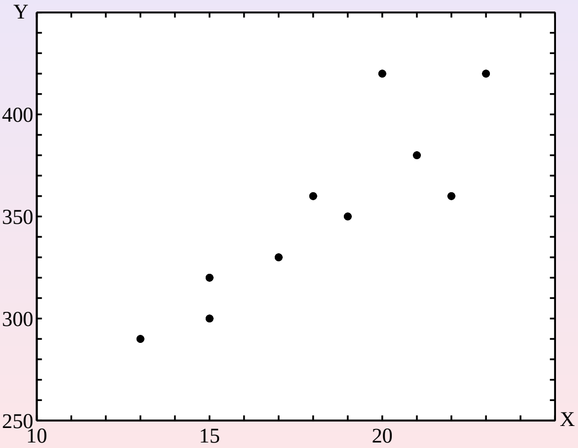


Figure: Nuage de points (x_i, y_i)

Le nuage statistique suggère un ajustement linéaire (les points sont presque alignés).

On va déterminer la droite de régression (D) de Y en fonction de X par la méthode des moindres carrées.

Pour cela, on complète le tableau suivant :

x_i	y_i	x_i^2	y_i^2	$x_i y_i$
15	300	225	90000	4500
21	380	441	144400	7980
17	330	289	108900	5610
22	360	484	129600	7920
20	420	400	176400	8400
13	290	169	81400	3770
19	350	361	122500	6650
15	320	225	102400	4800
18	360	324	129600	6480
23	420	529	176400	9660
183	3530	3447	1264300	65770

On obtient donc :

$$\overline{X} = \frac{183}{10} = 18.3$$

$$\overline{Y} = \frac{3530}{10} = 353$$

$$V(X) = \frac{3447}{10} - (18.3)^2 = 9.81$$

$$V(Y) = \frac{1264300}{10} - (353)^2 = 1821$$

$$Cov(X, Y) = \frac{65770}{10} - 18.3 \times 353 = 117.1$$

Donc les valeurs de a et b sont données par :

$$a = \frac{117.1}{9.81} = 11.9368$$

$$b = 353 - 11.9368 \times 18.3 = 134.5566$$

D'où l'équation de la droite d'ajustement de Y en fonction de X :

$$(D) : Y = 11.94X + 134.56$$

De même, on obtient l'équation de la droite d'ajustement de X en fonction de Y :

$$(D') : X = 0.064Y - 4.400$$

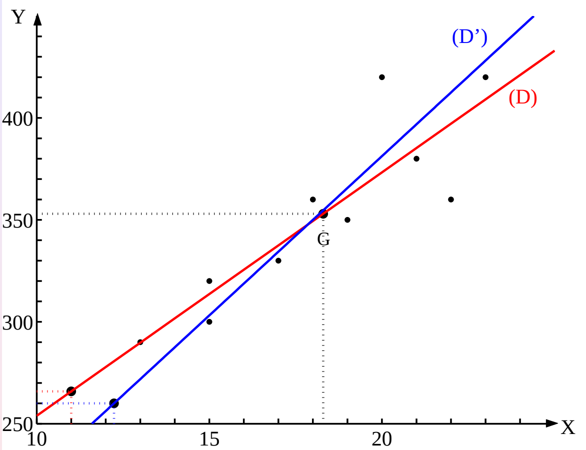


Figure: Droites de régression.

III.3- Coefficient de corrélation linéaire

Lorsque l'on a estimé la droite de régression, on doit se demander si cette estimation est de bonne qualité.

Pour cela on définit le coefficient de corrélation qui permet de mesurer l'intensité de dépendance linéaire entre les deux variables X et Y.

Définition

Le coefficient de corrélation linéaire d'un couple de variables statistiques quantitatives (X, Y), *noté* $r(X, Y)$, *est défini par :*

$$r(X,Y) = \frac{Cov(X,Y)}{\sigma(X)\sigma(Y)}$$

Propriétés

- r, $Cov(X, Y)$, a et a' sont tous de même signe.
- $-1 \leq r \leq 1$
- $r^2 = aa'$

Remarques

- Le coefficient de corrélation nous renseigne sur l'intensité de la dépendance linéaire entre les deux variables X et Y. L'ajustement est d'autant meilleur que $|r|$ est proche de 1 et il est d'autant plus mauvais que $|r|$ est proche de 0.
- $r = \pm 1 \Leftrightarrow$ relation linéaire parfaite du type $Y = aX + b$: corrélation linéaire parfaite
- Une corrélation linéaire parfaite ($r = \pm 1$) se traduit graphiquement par : les droites D et D' sont confondues.

- $r = 0 \Leftrightarrow$ aucune dépendance linéaire entre les deux variables.
- Un coefficient de corrélation nul ne signifie pas l'indépendance entre les deux variables. Il peut exister une relation non linéaire entre elles.

Exemple Reprenons l'exemple **2.**

On a

$$Cov(X, Y) = 117.1$$

$$V(X) = 9.81$$

$$V(Y) = 1821$$

d'où :

$$r = \frac{117.1}{\sqrt{9.81}\sqrt{1821}} = 0.876$$

La valeur de r est proche de 1 (> 0.7).
On a donc une bonne corrélation linéaire entre les deux variables X et Y.

Donc, on peut espérer de bonne estimation par la droite de régression.

Par exemple, le diamètre X d'un arbre de hauteur $Y = 450$ cm peut être estimé par la droite de régression (D') d'équation : $X = 0.064Y - 4.400$. Soit :

$$X = 0.064 \times 450 - 4.4 = 24.4 \; cm$$

III.4- Exemples d'ajustement non linéaire

Ajustement exponentiel

On suppose que les variables X et Y sont reliées par une relation du type :

$$Y = k \cdot \exp(aX)$$

En prenant le logarithme de cette expression, on obtient :

$$\ln Y = aX + \ln k$$

En posant $Z = \ln Y$ et $b = \ln k$, on obtient l'équation d'une droite $Z = aX + b$.

Ainsi, on se ramène à un problème d'ajustement linéaire.

Exemple3

Etude de l'évolution d'une population en extinction, composée initialement de 200 individus (relevés mensuels effectués pendant 8 mois) :

Mois (x_i)	0	1	2	3	4	5	6	7	8
Nbre d'individus (y_i)	200	121	74	45	27	16	10	6	4

Nuage statistique :

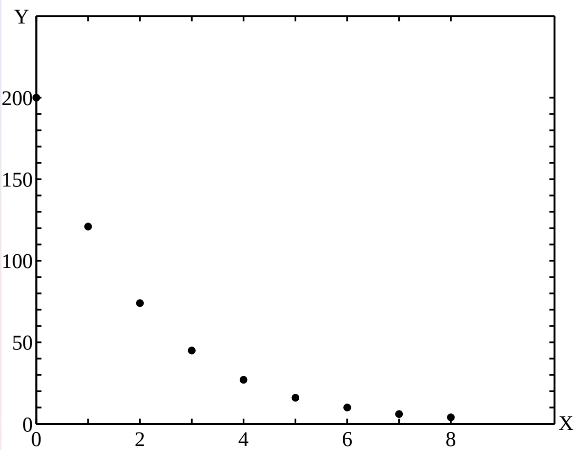


Figure: Nuage de points (x_i, y_j).

L'allure du nuage de point (x_i, y_i) montre qu'il existe entre les variables X et Y une relation non linéaire (exponentiel).

On va ajuster la distribution de (X, Y) par une fonction exponentielle : $Y = k\exp(aX)$.

Pour cela, on pose le changement de variables $Z = \ln Y$.

On obtient un nouveau tableau et un nouveau nuage de points (x_i, z_i)

x_i	0	1	2	3	4	5	6	7	8
z_i	5.3	4.8	4.3	3.8	3.3	2.8	2.3	1.8	1.4

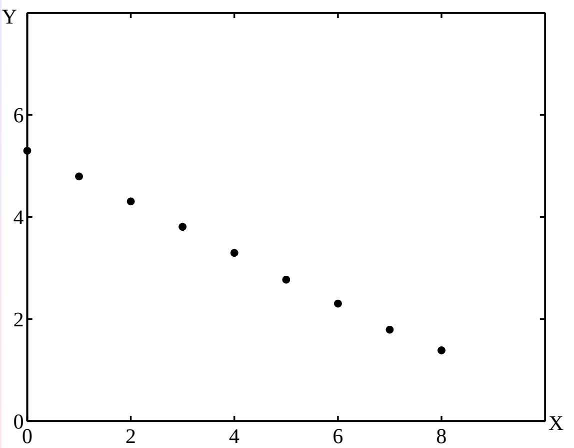


Figure: Nuage de points (x_i, z_j).

Le nuage de points (x_i, z_i) suggère maintenant un ajustement linéaire.

La droite de régression de Z en fonction de X par la méthode des moindres carrées a pour équation :

$$Z = -0.49X + 5.3$$

On déduit une modélisation de Y en fonction de X donnée par :

$$Y = 200\exp(-0.49X)$$

Ajustement puissance

On suppose que les variables X et Y sont reliées par une relation du type :

$$Y = k(X)^a$$

En prenant le logarithme de cette expression, on obtient :

$$\ln Y = a \ln X + \ln k$$

En posant $Z = \ln Y$, $T = \ln X$ et $b = \ln k$, on obtient l'équation d'une droite $Z = aT + b$.

Ainsi, on se ramène à un problème d'ajustement linéaire.

Exemple4

Considérons la série statistique suivante :

x_i	1	2	3	4	5	6	7	8	9	10
y_i	2	5	9	13	17	22	26	32	37	42

Nuage statistique :

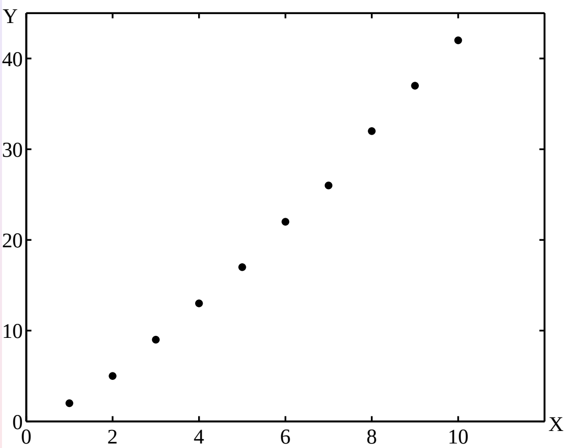


Figure: Nuage de points (x_i, y_j).

L'allure du nuage de point (x_i, y_i) montre que l'ajustement des variables (X, Y) ne peut pas être linéaire.

Nous allons alors ajuster la distribution de (X, Y) par une fonction puissance de la forme : $Y = k(X)^a$.

Pour cela, on pose le changement de variables $Z = \ln Y$ et $T = \ln X$. On obtient un nouveau tableau et un nouveau nuage de points (t_i, z_i) :

t_i	0	0.693	1.099	1.386	1.609	1.792	1.946	2.079
z_i	0.693	1.609	2.197	2.565	2.833	3.091	3.258	3.466

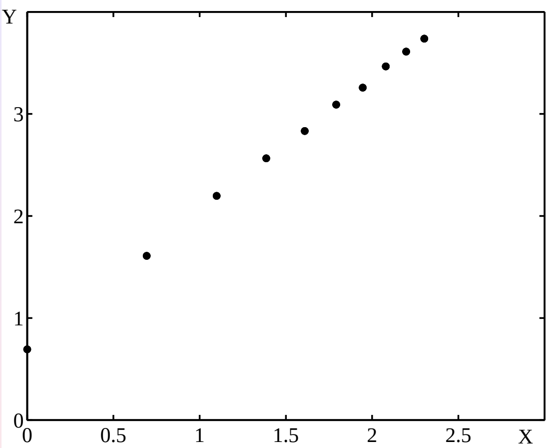


Figure: Nuage de points (t_i, z_j).

Le nuage de points (t_i, z_i) suggère maintenant un ajustement linéaire.

La droite de régression de Z en fonction de T par la méthode des moindres carrées a pour équation :

$$Z = 1.322\ T + 0.307$$

La modélisation de Y en fonction de X est donnée par :

$$Y = 2.031(X)^{1.322}$$

Chapitre 3

Calcul des Probabilités

Introduction

Le calcul des probabilités intervient dans l'étude des phénomènes aléatoires pour lesquels le résultat est imprévisible.

Le calcul des probabilités intervient dans l'étude des phénomènes aléatoires pour lesquels le résultat est imprévisible.

Le but principal de ce chapitre est de se familiariser avec le raisonnement probabiliste.

Sommaire

I- Notions de base

I.1 Expérience (ou épreuve) aléatoire

Définition

On appelle expérience (ou épreuve) aléatoire toute expérience dont le résultat est imprévisible, mais dont l'ensemble de tous les résultats possibles est connu à l'avance.

Chacun des résultats possible a une certaine chance d'être réalisé.

Exemples :

- Lancer un dé.
- Choisir une boule dans une urne.
- Effectuer une prise de sang sur un individu.

I.2 Ensemble fondamental (ou univers)

Définition

L'ensemble de tous les résultats possibles d'une expérience aléatoire est appelé univers (ou ensemble fondamental) associé à cette expérience. Il est noté Ω.

Exemple :

- Lancer un dé non truqué : $\Omega = \{1, 2, 3, 4, 5, 6\}$
- Déterminer le groupe sanguin d'un individu : $\Omega = \{A, B, AB, O\}$
- Lancer deux dés : $\Omega = \{(i, j), \;\; 1 \leq i \leq 6; \;\; 1 \leq j \leq 6\}$

I.3 Evénements

Soit Ω un univers associé à une épreuve aléatoire.

Définition

On appelle événement toute partie de Ω, dont on sait dire, à l'issue de l'épreuve, s'il est réalisé ou non.
Un événement (noté A) est caractérisé par le sous-ensemble $A \subset \Omega$ des résultats $\omega_i \in \Omega$ qui le réalisent.

Exemple

- On lance un dé non truqué.
  Soit l'événement $A =$ "obtenir un point pair". On a $A = \{2, 4, 6\}$.
  A l'issue de l'épreuve, on peut affirmer sans ambiguité si A est réalisé ou non.
- On lance deux dés non truqués.
  Soit l'événement $A =$ "la somme des points est supérieure à 10". On a $A = \{(5, 6), (6, 5), (6, 6)\}$
- On mesure la glycémie d'un individu.
  Soit l'événement $A =$ "la glycémie vaut $\sqrt{2}$".
  $A = \{\sqrt{2}\}$ est bien une partie de Ω mais il est impossible de savoir s'il est réalisé ou pas.

Remarques :

- On dit que l'événement A est réalisé si le résultat de l'épreuve aléatoire est un élément du sous-ensemble A.
- Toute partie de Ω n'est pas forcément un événement.
- L'ensemble de tous les événements associés à une épreuve aléatoire, noté $\mathcal{E}(\Omega)$, vérifie :

$$\mathcal{E}(\Omega) \subset \mathcal{P}(\Omega)$$

  où $\mathcal{P}(\Omega)$ est l'ensemble des parties de Ω.
- Si Ω est fini, on a :

$$\mathcal{E}(\Omega) = \mathcal{P}(\Omega)$$

I.4 Evénements remarquables

a) Evénement élémentaire

Un événement est dit élémentaire s'il est un événement constitué d'un seul élément $\omega \in \Omega$. On le note $\{\omega\}$ et on impose que ce soit un événement :

$$\{\omega\} \in \mathcal{E}(\Omega)$$

b) Evénement impossible

On appelle événement impossible, l'événement qui ne peut être réalisé, quelle que soit l'issue de l'épreuve.
On le note $\emptyset$ et on impose que ce soit un événement :

$$\emptyset \in \mathcal{E}(\Omega)$$

I.4 Evénements remarquables

c) Evénement certain

On appelle événement certain, l'événement qui est toujours réalisé quelle que soit l'issue de l'épreuve.
On le note Ω et on impose que ce soit un événement :

$$\Omega \in \mathcal{E}(\Omega)$$

I.4 Evénements remarquables

d) Evénement contraire

On appelle événement contraire d'un événement A, l'événement qui est réalisé si et seulement si A ne l'est pas. On le note $\overline{A}$ et on impose que ce soit un événement :

$$\forall\, A \in \mathcal{E}(\Omega), \quad \overline{A} \in \mathcal{E}(\Omega)$$

Diagramme :

I.4 Evénements remarquables

e) Intersection de deux événements

On appelle intersection de deux événements A et B, l'événement qui est réalisé si et seulement si A et B sont réalisés simultanément.
On le note $A \cap B$ et on impose que ce soit encore un événement :

$$\forall\ A,\ B \in \mathcal{E}(\Omega), \quad A \cap B \in \mathcal{E}(\Omega)$$

Diagramme :

I.4 Evénements remarquables

f) Union de deux événements

On appelle union de deux événements A et B, l'événement qui est réalisé si et seulement si l'un au moins des événements A ou B est réalisé.
On le note $A \cup B$ et on impose que ce soit encore un événement :

$$\forall\ A,\ B \in \mathcal{E}(\Omega), \quad A \cup B \in \mathcal{E}(\Omega)$$

Diagramme :

On peut généraliser les deux dernières définitions à n événements.

I.4 Evénements remarquables

g) Implication

On dit que l'événement A implique l'événement B lorsque la réalisation de A entraîne la la réalisation de B.
On note : $A \subset B$

I.4 Evénements remarquables

h) Evénements incompatibles

Deux événements A et B sont dits incompatibles s'ils ne peuvent être réalisés simultanément.
Autrement dit, A et B sont incompatibles si et seulement si

$$A \cap B = \emptyset$$

Diagramme :

Exemple

On jette un dé et on considère les événements :
$A =$ "obtenir un point pair".
$B =$ "obtenir un point impair".
A et B sont incompatibles.

Remarques

- Deux événements contraires A et $\overline{A}$ sont incompatibles. De plus, leur union est égale à Ω. On dit qu'ils forment une partition de Ω (ou un système complet).
- Inversement, deux événements incompatibles ne sont pas nécessairement contraires.

Remarque

On vient de définir sur les événements d'une expérience aléatoire des opérations logiques correspondantes aux opérations définies sur les parties d'un ensemble.

Ainsi, on établit une correspondance entre le langage des événements (language probabiliste) et le langage des ensembles (langage ensembliste).

Le tableau suivant résume la correspondance de vocabulaire et de notations :

Voc. probabiliste	Voc. ensembliste	Notations
Ev. A	ss-ensemble A de Ω	$A \subset \Omega$
Ev. impossible	ensemble vide	$\emptyset$
Ev. certain	ens. de ts les résultats	Ω
ω réalise A	ω appartient à A	$\omega \in A$
Ev. A ou B	union de A et B	$A \cup B$
Ev. A et B	intersection de A et B	$A \cap B$
Ev. contraire de A	complémentaire de A	$\overline{A}$
A et B incompatibles	A et B sont disjoints	$A \cap B = \varnothing$
A implique B ($A \Rightarrow B$)	A inclu dans B	$A \subset B$

Exemple

On lance un dé non truqué : $\Omega = \{1, 2, 3, 4, 5, 6\}$; ($\Omega$ fini)

On a : $\mathcal{E}(\Omega) = \mathcal{P}(\Omega)$; ($Card(\mathcal{P}(\Omega)) = 2^{Card(\Omega)}$).

Soient les événements :

$A =$ "obtenir un point ≤ 4".

$B =$ "obtenir un point ≥ 3".

On a :

$A \cap B$=$\{3, 4\}$

$A \cup B$=Ω

$\overline{A} = \{5, 6\}$

$\overline{B} = \{1, 2\}$

$\{2\} \Rightarrow A$

Propriétés

On retrouve pour les événements des propriétés analogues à celles connues sur les ensembles :

Si A, B et C sont des événements, alors :

- $\overline{\overline{A}} = A$
- $\overline{\emptyset} = \Omega$ et $\overline{\Omega} = \emptyset$
- $\overline{A \cap B} = \overline{A} \cup \overline{B}$ et $\overline{A \cup B} = \overline{A} \cap \overline{B}$ (Lois de Morgan)
- $A \cap \overline{A} = \emptyset$ et $A \cup \overline{A} = \Omega$
- $A \cap A = A$ et $A \cup A = A$
- $A \cap \Omega = A$ et $A \cup \emptyset = A$
- $A \cap \emptyset = \emptyset$ et $A \cup \Omega = \Omega$

Propriétés

- $A \cap B = B \cap A \quad \text{et} \quad A \cup B = B \cup A$
- $A \cap (B \cap C) = (A \cap B) \cap C \quad \text{et} \quad A \cup (B \cup C) = (A \cup B) \cup C$
- $A \cap (B \cup C) = (A \cap B) \cup (A \cap C) \quad \text{et}$
  $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$

II- Espaces probabilisés

II.1- Tribu d'événements

Soit Ω un univers associé à une épreuve aléatoire.

Définition

On appelle tribu d'événements de Ω, notée $\mathcal{T}$, toute famille de parties de Ω (i.e. $\mathcal{T} \subset \mathcal{P}(\Omega)$) vérifiant les propriétés suivantes :

1. $\Omega \in \mathcal{T}$
2. *Si $A \in \mathcal{T}$, alors $\overline{A} \in \mathcal{T}$*
3. *Si $(A_i)_i$est une famille dénombrable d'éléments de $\mathcal{T}$, alors $\bigcup_i A_i \in \mathcal{T}$*

Tout élément de la tribu $\mathcal{T}$ est un événement.

II.2- Définition de la probabilité

Soit Ω un univers et $\mathcal{T}$ une tribu d'événements de Ω.

Définition

On appelle probabilité sur $(\Omega, \mathcal{T})$ toute application :

$$p : \mathcal{T} \to [0, 1]$$

vérifiant :

(i) $p(\Omega) = 1$

(ii) $\forall\, A, B \in \mathcal{T}$, si A et B sont incompatibles alors :

$$p(A \cup B) = p(A) + p(B)$$

.

Définition

On appelle espace probabilisé tout triplet $(\Omega, \mathcal{T}, p)$ où Ω désigne un univers, $\mathcal{T}$ une tribu d'événements de Ω et p une probabilité sur $(\Omega, \mathcal{T})$.

Remarques :

- L'idée est donc d'associer à chaque événement $A \in \mathcal{T}$ un nombre réel positif compris entre 0 et 1, noté $p(A)$, qui mesure la chance (ou la fréquence) qu'à l'événement A pour se réaliser.

- $p(A)$ est appelé probabilité de l'événement A.

- Si Ω est fini, on prend $\mathcal{T} = \mathcal{P}(\Omega)$

Propriétés d'une probabilité

Soit $(\Omega, \mathcal{T}, p)$ un espace probabilisé.

- si $A_1, A_2, \ldots, A_n$ sont n événements deux à deux incompatibles alors :

$$p(\bigcup_{k=1}^{n} A_k) = \sum_{k=1}^{n} p(A_k)$$

- Si A est un événement, alors :

$$p(\overline{A}) = 1 - p(A)$$

- si $A \subset B$ alors $p(A) \leq p(B)$

Propriétés d'une probabilité

- Si A et B sont deux événements quelconques, alors :

$$p(A \cup B) = p(A) + p(B) - p(A \cap B)$$

- si $A_1, A_2, \ldots, A_n$ sont n événements quelconques, alors :

$$p(\bigcup_{k=1}^{n} A_k) \leq \sum_{k=1}^{n} p(A_k)$$

- $$p(\bigcup_{k=1}^{n} A_k) = \text{Formule du crible (voir TD)}$$

II.3- Univers dénombrable (fini ou infini)

Définition

Un univers Ω est dit dénombrable si on peut numéroter tous ses éléments.

exemple :

- Un univers fini est dénombrable ($\Omega = \{\omega_1, \omega_2, \dots, \omega_N\}$)

- $\Omega = \{\omega_1, \omega_2, \dots, \omega_k, \dots\}$ est un univers infini dénombrable.

- $\Omega = [0;\ 20]$ est un univers infini non dénombrable.

II.3- Univers dénombrable (fini ou infini)

Théorème

Soit Ω un univers dénombrable (chaque élément ω_i définit un événement élementaire $\{\omega_i\}$).

Une probabilité sur Ω est complètement déterminée, de façon unique, par la donnée des probabilités des événements élémentaires $p(\{\omega_i\})$, grâce à la formule :

$$\forall\ A \in \mathcal{E}(\Omega), \quad p(A) = \sum_{\omega_i \in A} p(\{\omega_i\})$$

On a :

$$\sum_{\omega_i \in \Omega} p(\{\omega_i\}) = 1$$

II.4- Cas particulier : Probabilité uniforme

Définition

Soit $\Omega = \{\omega_1, \omega_2, \ldots, \omega_N\}$ un univers fini (i.e. $(\mathcal{E}(\Omega) = \mathcal{P}(\Omega))$.

Une probabilité p sur Ω est dite uniforme si tous les événements élémentaires ont la même chance de se réaliser :

$$\forall\ 1 \leq i, j \leq N, \quad p(\{\omega_i\}) = p(\{\omega_j\})$$

On dit également qu'on a équiprobabilité.

Remarque

En général, on n'a pas à démontrer qu'on a équiprobabilité. C'est plutôt l'énoncé qui le suggère.

II.4- Cas particulier : Probabilité uniforme

Proposition

Soient Ω un univers fini de cardinal N et p une probabilité uniforme sur Ω. Alors on a :

- *p est unique,*
- $\forall\ 1 \leq i \leq N, \quad p(\{\omega_i\}) = \frac{1}{N}$
- $\forall\ A \in \mathcal{P}(\Omega)$,

$$p(A) = \frac{Card(A)}{Card(\Omega)} = \frac{\text{Nombre de cas favorables}}{\text{Nombre de cas possibles}}$$

Exemple

On lance un dé non truqué, donc on a équiprobabilité et Ω est fini :

$$\forall\, \omega_i \in \Omega, \quad p(\{\omega_i\}) = \frac{1}{6}$$

Soit l'événement $A =$ "obtenir un point pair". On a :

$$p(A) =$$

Exemple

On lance un dé non truqué, donc on a équiprobabilité et Ω est fini :

$$\forall\, \omega_i \in \Omega, \quad p(\{\omega_i\}) = \frac{1}{6}$$

Soit l'événement $A =$ "obtenir un point pair". On a :

$$p(A) = p(\{2,4,6\}) = \frac{3}{6} = 0.5$$

Exemple

On lance deux dés non truqués. On a équiprobabilité et Ω est fini.
Soit l'événement $A =$ "obtenir au moins un point pair". On a :

$$p(A) =$$

Exemple

On lance deux dés non truqués. On a équiprobabilité et Ω est fini.
Soit l'événement $A =$ "obtenir au moins un point pair". On a :

$$p(A) = \frac{27}{36}$$

Exemple

On lance deux dés non truqués, un noir et un blanc.
Soit les événements :
$N =$ "le numéro le plus élevé apparaît sur le dé noir"
$B =$ "le numéro le plus élevé apparaît sur le dé blanc"
$E =$ "les deux dés ont le même numéro"

On a :

$$p(E) =$$

Exemple

On lance deux dés non truqués, un noir et un blanc.
Soit les événements :
$N =$ "le numéro le plus élevé apparaît sur le dé noir"
$B =$ "le numéro le plus élevé apparaît sur le dé blanc"
$E =$ "les deux dés ont le même numéro"

On a :

$$p(E) = \frac{1}{6}$$

$$p(N) = p(B) =$$

Exemple

On lance deux dés non truqués, un noir et un blanc.
Soit les événements :
$N =$ "le numéro le plus élevé apparaît sur le dé noir"
$B =$ "le numéro le plus élevé apparaît sur le dé blanc"
$E =$ "les deux dés ont le même numéro"

On a :

$$p(E) = \frac{1}{6}$$

$$p(N) = p(B) = \frac{5}{12}$$

Exemple

Une urne contient 3 boules blanches, 4 boules rouges et 6 boules noires.
On tire une boule au hasard, donc on équiprobabilité.

Soit les événements :
$B =$ "obtenir une boule blanche"
$R =$ "obtenir une boule rouge"
$N =$ "obtenir une boule noire".

On a :

Exemple

Une urne contient 3 boules blanches, 4 boules rouges et 6 boules noires.
On tire une boule au hasard, donc on équiprobabilité.

Soit les événements :
$B =$ "obtenir une boule blanche"
$R =$ "obtenir une boule rouge"
$N =$ "obtenir une boule noire".

On a :
$$p(B) = \frac{3}{13};\ p(R) = \frac{4}{13};\ p(N) = \frac{6}{13}.$$

Remarque1

Dans le cas d'une probabilité uniforme, Les dénombrements des cas favorables et des cas possibles peuvent nécessiter des calculs d'analyse combinatoire (dénombrement), rappelés au dernier paragraphe de ce chapitre.

Exemple

Dans une urne contenant 9 boules numérotées de 1 à 9 et 26 boules numérotées de A à Z, on tire au hasard 6 boules pour former un mot de passe.

Quelle est la probabilité d'obtenir un mot de passe contenant autant de chiffres que de lettres ?

Remarque2

Le cas particulier de probabilité uniforme ne se présente pas toujours. L'univers Ω peut être infini et même s'il est fini, les événements élémentaires peuvent ne pas avoir tous la même chance de se réaliser.

Exemple(Ω infini dénombrable)

On lance un dé non truqué jusqu'à l'apparition du point $n^\circ 6$.
On s'intéresse au nombre de lancers nécessaires.
On a Ω infini dénombrable :

$$\Omega =$$

Exemple(Ω infini dénombrable)

On lance un dé non truqué jusqu'à l'apparition du point $n^\circ 6$.
On s'intéresse au nombre de lancers nécessaires.
On a Ω infini dénombrable :

$$\Omega = \{1, 2, \ldots, k, \ldots\} = \mathbb{N}^*$$

On a :

$$\forall\, k \in \Omega, \quad p(k) =$$

Exemple(Ω infini dénombrable)

On lance un dé non truqué jusqu'à l'apparition du point $n^\circ 6$.
On s'intéresse au nombre de lancers nécessaires.
On a Ω infini dénombrable :

$$\Omega = \{1, 2, \dots, k, \dots\} = \mathbb{N}^*$$

On a :

$$\forall\, k \in \Omega, \quad p(k) = \left(\frac{5}{6}\right)^{k-1} \times \frac{1}{6}$$

vérifier que : $\displaystyle\sum_{k=1}^{+\infty} p(k) = 1$

Soit l'événement $A =$ "effectuer au moins deux lancers". On a :

$$p(A) =$$

Soit l'événement $A =$ "effectuer au moins deux lancers". On a :

$$p(A) = \frac{5}{6}$$

Exemple (Ω non dénombrable)

- On lance deux dés non truqués, un noir et un blanc. On s'intéresse à la distance séparant les deux dés. On a alors une infinité d'éventualités et on peut considérer que :

$$\Omega = [0, d]$$

Exemple (Ω non dénombrable)

- On lance deux dés non truqués, un noir et un blanc. On s'intéresse à la distance séparant les deux dés. On a alors une infinité d'éventualités et on peut considérer que :

$$\Omega = [0, d]$$

- On effectue une prise de sang sur un individu choisi au hasard. On s'intéresse à la mesure de la glycémie.
On a encore une infinité d'éventualités et on peut considérer que :

$$\Omega = [0, a]$$

III- Probabilités conditionnelles

Question

Le fait de savoir qu'un événement B est réalisé peut-il modifier la probabilité de réalisation d'un autre événement A ?

Exemple

La répartition d'un groupe d'étudiants selon la note du module de mathématique et le sexe est représentée dans le tableau suivant :

Sexe / Note de Math	Masculin	Féminin	Total
$[0,7[$	4	3	7
$[7,10[$	17	15	32
$[10,20[$	21	36	57
Total	42	54	96

Un étudiant est choisi au hasard.
On considè les événements suivants :
$A =$ "l'étudiant choisi est un garçon"
$B =$ "l'étudiant choisi a validé le module"

On a :
$p(A) =$
$p(B) =$
$p(A \cap B) =$

On suppose maintenant qu'on sait que l'étudiant choisi au hasard a validé le module.

La probabilité qu'il soit un garçon est alors :

On suppose maintenant qu'on sait que l'étudiant choisi au hasard a validé le module.

La probabilité qu'il soit un garçon est alors : $\dfrac{21}{57}$.

On suppose maintenant qu'on sait que l'étudiant choisi au hasard a validé le module.

La probabilité qu'il soit un garçon est alors : $\dfrac{21}{57}$.

Si on note cette probabilité $p_B(A)$, on a :

$$p_B(A) = \frac{p(A \cap B)}{p(B)}$$

On suppose maintenant qu'on sait que l'étudiant choisi au hasard a validé le module.

La probabilité qu'il soit un garçon est alors : $\frac{21}{57}$.

Si on note cette probabilité $p_B(A)$, on a :

$$p_B(A) = \frac{p(A \cap B)}{p(B)}$$

Ainsi, La réalisation d'un événement B peut modifier la probabilité de réalisation d'un événement A.

III.1- Définition et propriétés

Définition

Soit $(\Omega, \mathcal{T}, p)$ un espace probabilisé et soient A et B deux événements tel que $p(B) \neq 0$.
La probabilité de réalisation de l'événement A sachant que l'événement B est réalisé, notée $p_B(A)$, est définie par :

$$p_B(A) = \frac{p(A \cap B)}{p(B)}$$

On l'appelle probabilité conditionnelle et on lit probabilité de A sachant B.

On utilise aussi la notation $p(A|B)$.

III.1- Définition et propriétés

Proposition

Soit $(\Omega, \mathcal{T}, p)$ *un espace probabilisé et soit* B *un événement tel que* $p(B) \neq 0$*, alors :*
L'application $p_B : \mathcal{T} \to [0,\ 1]$*, définie par :*

$$\forall\ A \in \mathcal{T}, \quad p_B(A) = \frac{p(A \cap B)}{p(B)}$$

est une probabilité sur Ω*, appelée probabilité conditionnelle de* A *sachant* B*.*

Remarque

Toutes les Propriétés d'une probabilité p sont ainsi valable pour une probabilité conditionnelle p_B.

Exemple

Une urne contient 3 boules blanches et 2 boules noires.
On tire au hasard, successivement et sans remise 2 boules.

Quelle est la probabilité que les deux boules tirées soient noires ?

Exemple

Une urne contient 3 boules blanches et 2 boules noires.
On tire au hasard, successivement et sans remise 2 boules.

Quelle est la probabilité que les deux boules tirées soient noires ?

On a $p(N_1) = \dfrac{2}{5}$; $p(N_2/N_1) = \dfrac{1}{4}$

donc $p(N_1 \cap N_2) =$

Exemple

Une urne contient 3 boules blanches et 2 boules noires.
On tire au hasard, successivement et sans remise 2 boules.

Quelle est la probabilité que les deux boules tirées soient noires ?

On a $p(N_1) = \dfrac{2}{5}$; $p(N_2/N_1) = \dfrac{1}{4}$

donc $p(N_1 \cap N_2) = p(N_2/N_1) \times p(N_1) = \dfrac{1}{10}$.

Formule des probabilités composées :

Proposition (Formule des probabilités composées)

Soient $A_1, A_2, \dots, A_n$ n événements tels que $p(A_1 \cap A_2 \cap \dots \cap A_n) \neq 0$. Alors :

$$p(A_1 \cap A_2 \cap \dots \cap A_n) =$$

$$p(A_1)p(A_2|A_1)p(A_3|A_2 \cap A_1)\dots p(A_n|A_1 \cap A_2 \cap \dots \cap A_{n-1})$$

Formule des probabilités composées :

Proposition (Formule des probabilités composées)

Soient $A_1, A_2, \dots, A_n$ n événements tels que $p(A_1 \cap A_2 \cap \dots \cap A_n) \neq 0$. Alors :

$$p(A_1 \cap A_2 \cap \dots \cap A_n) =$$

$$p(A_1)p(A_2|A_1)p(A_3|A_2 \cap A_1)\dots p(A_n|A_1 \cap A_2 \cap \dots \cap A_{n-1})$$

Exemple

Une urne contient 7 boules noires et 3 boules blanches. On tire successivement sans remise 3 boules.
Quelle est la probabilité de tirer 3 boules noires ?

Système complet d'événements

Définition

On dit que n événements $A_1, A_2, \ldots, A_n$ forment un système complet d'événements de Ω si :

1. Ils sont deux èa deux incompatibles :
   $\forall\ i \neq j, \quad A_i \cap A_j = \emptyset$
2. Leur réunion est l'événement certain :
   $A_1 \cup A_2 \cup \ldots \cup A_n = \Omega$

On dit aussi que les ensembles $A_1, A_2, \ldots, A_n$ forment une partition de Ω.

Formule des probabilités totales :

Proposition (Formule des probabilités totales)

Si $\{A_1, A_2, \ldots, A_n\}$ est un sysème complet d'événements (de probabilités non nulles), alors quel que soit l'événement B, on a:

$$p(B) = \sum_{i=1}^{n} p_{A_i}(B) \times p(A_i) = \sum_{i=1}^{n} p(A_i \cap B)$$

Exemple

On considère un groupe formé de 40 hommes et 60 femmes. Dans ce groupe, 50% des hommes sont fumeurs et 30% des femmes sont fumeuses.

On choisit un individu au hasard. Quelle est la probabilité qu'il soit un fumeur ?

Exemple

On considère un groupe formé de 40 hommes et 60 femmes. Dans ce groupe, 50% des hommes sont fumeurs et 30% des femmes sont fumeuses.

On choisit un individu au hasard. Quelle est la probabilité qu'il soit un fumeur ?

Quelle est la probabilité qu'il soit un homme, sachant qu'il est fumeur ?

Formule de Bayes :

Corollaire (Formule de Bayes)

Si $\{A_1, A_2, \ldots, A_n\}$ est un sysème complet d'événements (de probabilités non nulles) et si B est un événement de probabilité non nulle, alors :

$$p_B(A_j) = \frac{p(A_j) \times p_{A_j}(B)}{\displaystyle\sum_{i=1}^{n} p_{A_i}(B) \times p(A_i)}$$

Exemple

Trois machines M_1, M_2 et M_3 de fabrication mécanique produisent respectivement 25%, 35% et 40% de la production totale.

5% des pièces produites par M_1 sont défectueuses
4% des pièces produites par M_2 sont défectueuses
3% des pièces produites par M_3 sont défectueuses

On choisi au hasard une pièce de la production.

1. Quelle est la probabilité que la pièce choisie soit fabriquée par M_1 et qu'elle soit défectueuse ?
2. Quelle est la probabilité que la pièce choisie soit défectueuse ?
3. Si la pièce choisie est défectueuse, quelle est la probabilité qu'elle soit fabriquée par M_1 ?

III.2- Evénements indépendants

Définition

Deux événements A et B d'un même espace probabilisé sont dits indépendants si et seulement si la réalisation de l'un n'a aucune influence sur la probabilité de réalisation de l'autre. On a alors :

$$p(A/B) = p(A)$$

et

$$p(B/A) = p(B)$$

Corollaire

Deux événements A et B sont dits indépendants si et seulement si :

$$p(A \cap B) = p(A) \times p(B)$$

Exemple

On lance un dé deux fois. Soit les événements :

$A =$ "obtenir un point pair au premier lancer"

$B =$ "obtenir le point 1 ou 4 au deuxième lancer"

Les événements A et B sont-ils indépendants ?

exemple

On considère une famille et les événements :

$A =$ "la famille a des enfants de sexes différents"

$B =$ "la famille a au plus une fille"

Les événements A et B sont-ils indépendants dans les deux cas suivants : ?

- 1er cas : La famille est composée de 2 enfants.
- 2ème cas : La famille est composée de 3 enfants.

Remarque

- Il ne faut pas confondre événements indépendants et événements incompatibles :
  indépendants $\Rightarrow p(A \cap B) = p(A) \times p(B)$
  incompatibles $\Rightarrow p(A \cap B) = 0$
- Dans une suite de tirages successifs avec remise, les tirages sont indépendants.
- Dans une suite de tirages successifs sans remise, les tirages ne sont pas indépendants.
- si A et B sont indépendants, il en est de même pour les paires d'événements A et $\overline{B}$; $\overline{A}$ et B; $\overline{A}$ et $\overline{B}$.

exemple

On effectue une suite de tirages successifs dans une urne qui contient n boules dont r boules rouges.
Quelle est la probabilité pour que la première boule rouge apparaisse au troisième tirage ?

Généralisation à *n* evénements

Définition

n événements $A_1, A_2, \ldots, A_n$ d'un même espace probabilisé sont dits indépendants dans leurs ensemble si pour toute famille de p ($p \geq 2$) événements $A_{i_1}, A_{i_2}, \ldots, A_{i_p}$ choisis parmi les événements $A_1, A_2, \ldots, A_n$, on a :

$$p(A_{i_1} \cap A_{i_2} \cap \ldots \cap A_{i_p}) = p(A_{i_1}) \times p(A_{i_2}) \times \ldots \times p(A_{i_p})$$

Remarque

n événements $A_1, A_2, \ldots, A_n$ peuvent être indépendants deux à deux sans être indépendants dans leurs ensemble.

Exemple

On jette deux dés et on considère les événements suivants :

$A_1 =$ "obtenir un point pair par le 1er dé"

$A_2 =$ "obtenir un point pair par le 2ème dé"

$A_3 =$ "la somme des points obtenus par les deux dés est paire"

IV- Dénombrement

On se place dans le cas Ω fini et équiprobabilité.
On a alors pour tout événement A :

$$p(A) = \frac{\text{Nombre de cas favorables}}{\text{Nombre de cas possibles}} = \frac{Card(A)}{Card(\Omega)}$$

Remarque

Le dénombrement des cas favorables et des cas possibles peut nécessiter des calculs d'analyse combinatoire que nous allons rappeler dans ce dernier paragraphe.

IV.1- Cardinal

Définition

Soit E un ensemble fini. On appelle cardinal de E, noté $Card(E)$, le nombre d'éléments de E.

Exemple

Dans un groupe de 85 étudiants, 62 ont validé le module de Math, 54 ont validé le module de Physique et 20 n'ont validé aucun module.
Combien d'étudiants ont validé les deux modules ?

Propriétés :

- Soit $A \subset E$. On a : $Card(\overline{A}) = Card(E) - Card(A)$
- Si $A_1, A_2, ..., A_n$ constituent une partition de E, alors :

$$Card(E) = Card(A_1) + Card(A_2) + ... + Card(A_n)$$

- Soient E_1 et E_2 deux ensembles finis. On a :

$$Card(E_1 \times E_2) = Card(E_1) \times Card(E_2)$$

- Soient A et B deux parties de E. On a :

$$Card(A \cup B) = Card(A) + Card(B) - Card(A \cap B)$$

Propriétés :

- Soient $A_1, A_2, ..., A_n$, n parties de E. On a :

$$Card(\bigcup_{i=1}^{n} A_i) =$$

$$\sum_{k=1}^{n}\left((-1)^{k+1}\sum_{1\leq i_1<i_2...<i_k\leq n} Card(A_{i_1}\cap A_{i_2}\cap\ldots\cap A_{i_k})\right)$$

IV.2- Principe multiplicatif

Exemple

Un code comporte deux lettres suivies de trois chiffres non nuls. Combien peut-on former de codes distincts ?

IV.2- Principe multiplicatif

Exemple

Un code comporte deux lettres suivies de trois chiffres non nuls. Combien peut-on former de codes distincts ?

Principe

Si la réalisation d'un événement A se fait en k étapes présentant respectivement $n_1, n_2, \ldots, n_k$ possibilités, alors le nombre total de possibilités de réalisation de A est égale au produit : $n_1 \times n_2 \times \ldots \times n_k$.

IV.3- p-listes

Définition

Soit E un ensemble fini de cardinal n. Soit $p \in \mathbb{N}^*$.
On appelle liste de p éléments (ou p-liste) de E, tout p-uplet d'éléments de E.
C'est donc une suite de p éléments choisis successivement et avec remise parmi n éléments en tenant compte de l'ordre d'apparition.

Remarque

L'ensemble des p-listes de E est l'ensemble produit cartésien $E^p = E \times E \times \ldots \times E$, ($p$ fois).

Proposition

Le nombre des p-listes d'un ensemble E de cardinal n est égal à n^p.

Exemple

On tire successivement avec remise 4 boules, d'une urne qui contient 10 boules numérotées de 0 à 9.

Exemple

Au loto sportif, on coche l'une des trois cases 1, N ou 2 pour chacun des 15 matches sélectionnés.
Dénombrer le nombre de grilles possibles.

Exemple

Au loto sportif, on coche l'une des trois cases 1, N ou 2 pour chacun des 15 matches sélectionnés.
Dénombrer le nombre de grilles possibles.

Chaque grille est une 15-liste de l'ensemble $E = \{1, N, 2\}$. Il y a donc $3^{13} = 1594323$ grilles possibles.

Astuce

On fait appel au nombre de p-listes (n^p) lorsqu'on dénombre le nombre de façon de choisir au hasard, **successivement et avec remise**, p éléments parmi n (donc avec prise en compte de l'ordre et avec répétition).

IV.4- Arrangements

Définition

Soit E un ensemble fini de cardinal n. Soit $p \in \mathbb{N}^*$.
On appelle arrangement sans répétition de p éléments (ou p-arrangement) de E tout p-uplet d'éléments distincts de E.
C'est donc une suite de p éléments distincts choisis successivement et sans remise parmi n éléments en tenant compte de l'ordre d'apparition.

On note A_n^p le nombre d'arrangement sans répétition de p éléments parmi n.

Proposition

Le nombre d'arrangement sans répétition de p éléments choisis parmi n éléments de E, noté A_n^p, est égal à :

$$A_n^p = n(n-1)\dots(n-p+1) = \frac{n!}{(n-p)!}$$

Remarque

On doit avoir nécessairement $p \leq n$, sinon on pose $A_n^p = 0$.

Exemples

- Dans une urne contenant 10 boules numérotées de 0 à 9, on tire successivement sans remise 4 boules.
- Le nombre de façons de placer p objets distincts dans n cases, (chaque case pouvant contenir au plus 1 objet).
- Le nombre de tiercé possibles dans l'ordre avec 20 chevaux.

Astuce

On fait appel au nombre d'arrangement A_n^p lorsqu'on dénombre le nombre de façon de choisir, au hasard **successivement et sans remise**, p éléments parmi n (donc avec prise en compte de l'ordre mais sans répétition).

Permutations

Définition

Soit E un ensemble fini de cardinal n.
On appelle permutation de E tout arrangement de n éléments de E.
C'est donc une suite de n éléments distincts choisis successivement et sans remise parmi n éléments en tenant compte de l'ordre d'apparition.

Proposition

Le nombre de permutations de n éléments distincts d'un ensemble E est égal à $n!$

Exemple

- Le nombre de façons de placer 7 personnes sur 7 chaises.
- 4 livres de math, 3 livres de physique et 5 livres de chimie sont à placer dans 12 cases. On veut placer les livres de chaque catégorie les uns à coté des autres. Combien a-t-on de dispositions possibles ?

Permutations avec répétitions

Proposition

Supposons que certains des éléments de E sont identiques :
Supposons qu'on a k éléments distincts ($k \leq n$), chacun d'eux apparaissant $n_1, n_2, \ldots, n_k$ fois, avec $n_1 + n_2 + \cdots + n_k = n$, alors le nombre de permutations possibles est :

$$P_n(n_1, n_2, \ldots, n_k) = \frac{n!}{n_1!\, n_2! \ldots n_k!}$$

Exemple1

Une urne contient 2 boules rouges identiques, 3 boules blanches identiques et 4 boules noires identiques.
De combien de façons peut-on choisir les 9 boules ?

Exemple2

Combien de permutations peut-on effectuer avec les lettres du mot EVENEMENT ?

IV.5- Combinaisons

Définition

Soit E un ensemble fini de cardinal n. Soit $p \in \mathbf{N}^*$.
On appelle combinaison de p éléments (ou p-combinaison) de E, tout sous-ensemble constitué de p éléments de E.
C'est donc une famille de p éléments distincts choisis parmi n éléments sans tenir compte de l'ordre d'apparition.

On note C_n^p le nombre de combinaison de p éléments choisis parmi n.

Proposition

Le nombre de combinaison de p éléments choisis parmi n éléments de E, noté C_n^p, est égal à :

$$C_n^p = \frac{n!}{p!(n-p)!}$$

Remarque

On doit avoir nécessairement $p \leq n$, sinon on pose $C_n^p = 0$.

Exemples

- dans une urne contenant 49 boules numérotées de 1 à 49, on tire simultanément 6 boules.
- Le nombre de droites que l'on peut tracer à partir de 7 points
- Le nombre de façons de placer p objets indiscernables dans n cases (chaque case pouvant contenir au plus 1 objet)

- Dans un jeu de 32 cartes, on choisit 5 cartes au hasard (ces 5 cartes s'appellent une "main").
  1. Nombre de mains possibles :
  2. Nombre de mains qui contiennent exactement 3 as :
  3. Nombre de mains qui contiennent au moins 3 as :

Astuce

On fait appel au nombre de combinaison C_n^p lorsqu'on dénombre le nombre de façon de choisir, au hasard et **simultanément**, p éléments parmi n (donc sans répétition et sans prise en compte de l'ordre).

Propriétés

$$A_n^p = p!C_p^n$$

$$C_n^0 = C_n^n = 1$$

$$C_n^p = C_n^{n-p}$$

$$C_n^p = C_{n-1}^{p-1} + C_{n-1}^p$$

$$(a+b)^n = \sum_{p=0}^{n} C_n^p a^p b^{n-p}$$

$$\sum_{p=0}^{n} C_n^p = 2^n$$

Résumé des astuces de choix des méthodes de dénombrement

Méthode	Tirages	Ordre	Répétitions	Dénombrement
p-istes	S. A. R.	Oui	Oui	n^p
p-arrangements	S. S. R.	Oui	Non	A_n^p
p-combinaisons	Simultané	Non	Non	C_n^p

Chapitre 4

Variables aléatoires

Introduction

Le but de ce chapitre et du chapitre suivant est d'introduire des outils de calcul des probabilités avec les variables aléatoires.

Une variable aléatoire consiste à associer une valeur numérique à chaque élément de Ω.

Exemple

On jette deux dés.
A chaque élément $(i, j) \in \Omega$, on associe la somme $i + j$.
On définit ainsi une application $X : \Omega \to \mathbf{R}$.

L'ensemble des valeurs prises par cette application (noté $X(\Omega)$) est :

$$X(\Omega) =$$

Exemple

On jette deux dés.
A chaque élément $(i, j) \in \Omega$, on associe la somme $i + j$.
On définit ainsi une application $X : \Omega \to \mathbf{R}$.

L'ensemble des valeurs prises par cette application (noté $X(\Omega)$) est :

$$X(\Omega) = \{2, 3, 4, \dots, 12\}$$

X est une variable aléatoire sur Ω.

I- Définition

I- Définition

Définition

Soit $(\Omega, \mathcal{T}, p)$ *un espace probabilisé.*
On appelle variable aléatoire définie sur Ω *toute application* $X : \Omega \to \mathbf{R}$ *telle que pour tout* $x \in \mathbf{R}$, *l'ensemble* $A = \{\omega \in \Omega \ / \ X(\omega) \leq x\}$ *soit un événement (ie.* $A \in \mathcal{T}$*).*

Si la tribu $\mathcal{T}$ est égale à $\mathcal{P}(\Omega)$, toute application de Ω dans $\mathbf{R}$ est une variable aléatoire.

Remarque :

De façon analogue aux caractères statistiques, on distinguera deux cas : les variables aléatoires discrètes et les variables aléatoires continues.

- Une variable aléatoire qui ne prend qu'un nombre fini ou dénombrable de valeurs ($X(\Omega)$ dénombrable) est qualifiée de discrète.
- Une variable aléatoire qui prend une infinité non dénombrable de valeurs ($X(\Omega)$ non dénombrable) est dite continue.

Dans la suite de ce chapitre, on ne considère que des variables aléatoires discrètes ($X(\Omega)$ dénombrable).

On note $X(\Omega) = \{x_1, x_2, ...\}$ l'ensemble des valeurs prises par la variable aléatoire X.

Notations :

Soit $X : \Omega \to \mathbf{R}$ une variable aléatoire discrète.

- Pour chaque valeur $x_i \in X(\Omega)$, l'ensemble $X^{-1}(x_i) = \{\omega \in \Omega / X(\omega) = x_i\}$ est un événement que l'on note également $(X = x_i)$.
- Soit $I \subset R$. L'ensemble $X^{-1}(I) = \{\omega \in \Omega,\ X(\omega) \in I\}$ est un événement. On le note $(X \in I)$.
- On utilise aussi la notation : $(X \leq a)$, $(X \geq a)$, $(a \leq X \leq b)$, ...

Exemples

- On lance trois fois une pièce de monnaie.
  Soit $X =$ nombre de piles obtenues.
  X est une variable aléatoire discrète.
  On a $X(\Omega) = \{0, 1, 2, 3\}$.

  $(X = 1)$ est l'événement $\{PFF, FPF, FFP\}$.

- On jette deux dés.
  Soit $X =$ la somme des points obtenus.
  X est une variable aléatoire discrète.

  On a $X(\Omega) = \{2, 3, \ldots, 12\}$.

  $(X = 3)$ est l'événement $\{(1, 3), (2, 2), (3, 1)\}$.

II- Loi de probabilité d'une variable aléatoire discrète

II- Loi de probabilité d'une variable discrète

Définition

Soit $(\Omega, \mathcal{T}, p)$ un espace probabilisé et soit X une variable aléatoire discrète sur Ω.

On appelle loi de probabilité de X, l'application, notée p_X, définie de $X(\Omega)$ dans $[0,1]$ qui à chaque valeur $x_i \in X(\Omega)$ fait correspondre la probabilité $p_X(x_i) = p(X = x_i)$.

Remarque :

- Pour définir la loi de probabilité d'une variable aléatoire discrète X, il faut et il suffit de connaître les valeurs $p_X(x_i), \ \forall\, x_i \in X(\Omega)$.
- On a :

$$p_X(x_i) = p(X = x_i) = \sum_{(\omega \in \Omega / \ X(\omega) = x_i)} p(\{\omega\})$$

- Si $X(\Omega)$ est fini, on peut représenter la loi de probabilité de X sous forme d'un tableau.
- On doit avoir $\sum_{x_i \in X(\Omega)} p_X(x_i) = 1$.

Exemple

On jette deux dés.
A chaque élément $(i, j) \in \Omega$, on associe la somme $i + j$. On définit ainsi une v.a.d. X

$$X(\Omega) = \{2, 3, 4, \ldots, 12\}$$

Loi de probabilité de X :

Exemple

On jette deux dés.
A chaque élément $(i, j) \in \Omega$, on associe la somme $i + j$. On définit ainsi une v.a.d. X

$$X(\Omega) = \{2, 3, 4, \ldots, 12\}$$

Loi de probabilité de X :

x_i	2	3	4	5	6	7	8	9	10	11	12
$p(x_i)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

Remarque :

On peut généraliser la notion de l'événement $(X = x_i)$ et définir pour tout intervalle $I \subset \mathbb{R}$, l'événement : $(X \in I) = \{\omega \in \Omega / \ X(\omega) \in I\}$.

On a alors $p_X(I) = p(X \in I) = \sum_{x_i \in I} p(X = x_i)$.

III- Fonction de répartition

III- Fonction de répartition

Définition

Soit $(\Omega, \mathcal{T}, p)$ *un espace probabilisé et soit* X *une variable aléatoire discrète sur* Ω.

On appelle fonction de répartition de X, *l'application, notée* F_X, *définie de* $\mathbb{R}$ *vers* $[0,1]$ *par :*

$$\forall\, t \in \mathbb{R}, F_X(t) = p(X \leq t)$$

$$F_X(t) = \sum_{x_i \leq t} p(X = x_i)$$

Propriétés :

- F_X est croissante,
- F_X est constante par intervalle (fonction en escalier),
- F_X est discontinue en x_i (continue à droite),

- $p(X > a) = 1 - F_X(a)$,
- $p(a < X \leq b) = F_X(b) - F_X(a)$,
- $p(X = a) = F_X(a) - \lim_{t \to a^-} F_X(t)$
- Le saut de discontinuité au point x_i est égal à $p(x_i)$.

IV- Paramètres d'une variable aléatoire

IV.1- Espérance mathématique

Définition

Soit X une variable aléatoire discrète. On appelle espérance mathématique de X le nombre réel (si il existe), noté $E(X)$, défini par :

$$E(X) = \sum_{x_i \in X(\Omega)} x_i p(X = x_i)$$

Remarque :

Si $X(\Omega)$ est fini, $E(X)$ existe toujours.
Si $X(\Omega)$ est infini, $E(X)$ existe sous réserve de convergence.

Remarque :

On a aussi :

$$E(X) = \sum_{\omega \in \Omega} X(\omega) p(\{\omega\})$$

Remarque :

Soit X une variable aléatoire discrète et soit f une fonction réelle. Alors $f(X)$ est une variable aléatoire discrète et on a :

$$E(f(X)) = \sum_{x_i \in X(\Omega)} f(x_i) p(X = x_i)$$

En particulier si $f(x) = x^r$, On obtient

$$E(X^r) = \sum_{x_i \in X(\Omega)} x_i^r p(X = x_i)$$

$E(X^r)$ s'appelle le moment d'ordre r de X.

IV.2- Variance - écart-type

Définition

On appelle variance de X le nombre réel (si il existe), noté $V(X)$, défini par :

$$V(X) = \sum_{x_i \in X(\Omega)} (x_i - E(X))^2 \, p(X = x_i)$$

$$V(X) = E\left((X - E(X))^2\right)$$

Formule simplifiée : $V(X) = E(X^2) - (E(X))^2$

$$V(X) = \left(\sum_{x_i \in X(\Omega)} x_i^2 p(X = x_i) \right) - (E(X))^2$$

Définition

On appelle écart-type de X le nombre, noté $\sigma(X)$, défini par :

$$\sigma(X) = \sqrt{V(X)}$$

Exemple

On lance un dé et on suppose que l'on gagne 10 dh si on obtient la face 1; 5 dh si on obtient la face 2 ou 3; 2 dh si on obtient la face 4 et on perd 5 dh si on obtient la face 5 ou 6.

On note X la v.a. qui représente le gain.

On a la loi de probabilité de X :

x_i	-5	2	5	10
p_i	1/3	1/6	1/3	1/6

On a la loi de probabilité de X :

x_i	-5	2	5	10
p_i	1/3	1/6	1/3	1/6

L'espérance mathématique du gain est
$E(X) = -5 * 1/3 + 2 * 1/6 + 5 * 1/3 + 10 * 1/6 = 2$ dh.

La variance est donnée par :
$V(X) = 25 * 1/3 + 4 * 1/6 + 25 * 1/3 + 100 * 1/6 - 4 = 30.$

Propriétés :

Soient X et Y deux variables aléatoires définies sur un même espace probabilisé et a et b deux constantes réelles, on a :

1. $E(aX + b) = aE(X) + b$
2. $V(aX + b) = a^2 V(X)$
3. $E(X + Y) = E(X) + E(Y)$
4. Si X et Y sont indépendantes alors :
   $E(XY) = E(X)E(Y)$
   $V(X + Y) = V(X) + V(Y)$

V- Lois de probabilité discrètes usuelles

V.1- Loi uniforme

Définition

On dit qu'une variable aléatoire discrète X suit une loi uniforme sur $\{1,\ 2,\ \dots,\ n\}$ ssi

$$X(\Omega) = \{1,\ 2\ \dots,\ n\}$$

et

$$\forall k \in X(\Omega) \quad p(X = k) = \frac{1}{n}$$

On note $X \to \mathcal{U}([1..n])$.

Exemple

Dans une urne contenant n boules numérotées de 1 à n, on tire une boule au hasard.
Soit la variable aléatoire qui désigne le numéro de la boule tirée.

Proposition

Si X est une variable aléatoire discrète qui suit une loi uniforme sur $[1..n]$ alors on a :

$$E(X) = \frac{n+1}{2}$$

$$V(X) = \frac{n^2-1}{12}$$

V.2- Loi binomiale

Epreuve de Bernoulli

Définition

On appelle épreuve de Bernoulli de paramètre p toute épreuve aléatoire n'ayant que deux résultats possibles :
S appelé succès avec la probabilité p
$\overline{S}$ appelé échec avec la probabilité $q = 1 - p$.

Exemple

- On lance une pièce de monnaie, $p = q = 0.5$.

- On lance un dé et on considère $S =$ "obtenir la face 5 ou 6".

Variable et Loi de Bernoulli

Définition

Soit $\Omega = \{S, \overline{S}\}$ *l'univers d'une épreuve de Bernoulli de paramètre p.*
soit X la variable aléatoire désignant le nombre de succès.

X est appelée variable de Bernoulli et la loi de X est appelée loi de Bernoulli de paramètre p.

Proposition

Si X est une variable aléatoire qui suit une loi de Bernoulli de paramètre p alors on a :

$$X(\Omega) = \{0, 1\}$$

$$\forall k \in X(\Omega) \quad p(X = k) = p^k q^{1-k}$$

$$E(X) = p$$

$$V(X) = p(1 - p) = pq$$

Variable et Loi binomiale

Définition

Soit X la variable aléatoire désignant le nombre de succès obtenus en répètant n fois, de manières indépendantes, une épreuve de Bernoulli de paramètre p.

X est appelée variable binomiale et la loi de X est appelée loi binomiale de paramètres n et p.

On note $X \to \mathcal{B}(n, p)$.

Proposition

Une variable aléatoire X suit une loi binomiale de paramètres n et p si et seulement si :

$$X(\Omega) = \{0, 1, \ldots, n\}$$

et

$$\forall k \in X(\Omega), \quad p(X = k) = C_n^k p^k q^{n-k}$$

Proposition

Soit X une variable aléatoire qui suit une loi binomiale de paramètres n et p. Alors on a :

$$E(X) = np$$

$$V(X) = npq$$

Reconnaissance

On reconnait qu'une variable aléatoire X suit une loi binomiale de paramètres n et p si X représente le nombre de succès obtenus en répétant n fois, de manière indépendante, une épreuve de Bernoulli de paramètre p.

Exemple

On lance n fois un dé.

- Quelle est la probabilité d'obtenir k fois le chiffre 5 ?
- Combien de lancers faut-il effectuer pour que la probabilité d'obtenir au moins une fois le chiffre 5 soit supérieure à 0.9 ?
- On suppose $n = 12$. Calculer la moyenne d'apparition du chiffre 5.

Remarque

- On a : $\sum_{k=0}^{n} C_n^k p^k q^{n-k} = 1$

- L'expression de la loi binomiale est le terme général des coefficients du binôme de Newton

$$(p+q)^n = \sum_{k=0}^{n} C_n^k p^k q^{n-k}$$

d'où le nom de loi binomiale.

- Pour les grandes valeurs de n les calculs de la loi binomiale deviennent très difficiles. On peut, sous certaines conditions, utiliser des approximations avec d'autres lois (voir plus loin).

V.3- Loi de Poisson

Définition

On dit qu'une variable aléatoire X suit une loi de Poisson de paramètre $\lambda > 0$ si et seulement si :

$$X(\Omega) = \mathbb{N}$$

et

$$\forall k \in X(\Omega) = \mathbb{N}, \quad p(X = k) = \frac{\lambda^k e^{-\lambda}}{k!}$$

on note $X \to \mathcal{P}(\lambda)$.

Remarque

Une variable aléatoire X qui suit une loi de Poisson est une variable aléatoire dénombrable ($X(\Omega) = \mathbb{N}$).

Proposition

Soit X une variable aléatoire qui suit une loi de Poisson de paramètre λ. Alors on a :

$$E(X) = \lambda$$

$$V(X) = \lambda$$

Reconnaissance

On reconnait qu'une variable aléatoire X suit une loi de Poisson de paramètre λ si X représente le nombre de succès obtenus en répétant, de manière indépendante, une épreuve de Bernoulli pendant une période donnée durant laquelle la moyenne de succès est égale à λ.

La loi de Poisson modélise le nombre d'occurrences d'un phénomène par unité de temps (nombre de clients à un guichet par heure, nombre d'arrivées d'avions à un aéroport par jour, nombre d'accidents sur une autoroute par mois).

Exemple

On admet que le nombre moyen d'appels téléphoniques reçus par un standard, durant une heure est égal à 10 appels.
Donner la probabilité que le nombre d'appels reçus durant une période de 6 mn soit ≥ 4.

Soit X la variable aléatoire qui représente le nombre d'appels reçus pendant la période $T = 6mn = 0.1heure$.
On a $X \to \mathcal{P}(1)$, Donc

$$P(X \geq 4) = 1 - P(X < 4) =$$

Approximation d'une loi binomiale

Proposition

Soit X une variable aléatoire qui suit une loi binomiable de paramètres n et p.
Si n est assez grand ($n \geq 50$) et p petit de telle sorte que le produit np soit ≤ 5 ($p \leq 0.1$), alors on peut approximer la loi binomiale $\mathcal{B}(n, p)$ par une loi de Poisson de paramètre $\lambda = np$:

$$C_n^k p^k q^{n-k} \approx \frac{\lambda^k e^{-\lambda}}{k!}$$

V.4- Loi géométrique

On répète, de manière indépendante, une épreuve de Bernoulli de paramètre p jusqu'à l'obtention du premier succès.

Définition

Soit X la variable aléatoire désignant le nombre de répétitions effectuées.

La loi de probabilité de X est appelée loi géométrique (ou loi de Pascal) de paramètres p.

On note $X \to \mathcal{G}(p)$.

Proposition

Une variable aléatoire X suit une loi géométrique de paramètres p Si et seulement si :

$$X(\Omega) = \mathbb{N}^*$$

et

$$\forall k \in X(\Omega), \quad p(X = k) = pq^{k-1}$$

Proposition

Soit X une variable aléatoire qui suit une loi géométrique de paramètres p. Alors on a :

$$E(X) = \frac{1}{p}$$

$$V(X) = \frac{q}{p^2}$$

Reconnaissance

On reconnait qu'une variable aléatoire X suit une loi géométrique de paramètre p si X représente le nombre de répétitions indépendantes d'une épreuve de Bernoulli de paramètre p jusqu'à ce que le premier succès se réalise.

Remarque

Une variable aléatoire X qui suit une loi géométrique est une variable aléatoire dénombrable ($X(\Omega) = \mathbb{N}^*$).

Exemple

Une urne contient 5 boules blanches et 10 boules noires. On tire des boules au hasard et avec remise jusqu'à ce qu'on obtienne la première boule blanche (succès).
Quelle est la probabilité que la première boule blanche soit tirée après 4 tirages ?

Soit X la variable aléatoire qui représente le nombre de boules tirées jusqu'à l'obtention d'une boule blanche.

$$X \to \mathcal{G}(\frac{1}{3})$$

$$p(X = 4) = (\frac{2}{3})^3 \frac{1}{3}$$

V.5- Loi hypergéométrique

Définition

Soit $E = S \cup \overline{S}$ un ensemble formé de $N = (N_1 + N_2)$ éléments. On suppose S formé de N_1 éléments $\overline{S}$ formé de N_2 éléments. On choisit au hasard, simultanément ou successivement sans remise, n éléments dans l'ensemble E.
Soit X la variable aléatoire désignant le nombre d'éléments choisis appartenant à S (succès).

La loi de probabilité de X est appelée loi hypergéométrique de paramètres N_1, N_2 et n.

Proposition

Une variable aléatoire X suit une loi hypergéométrique de paramètres N_1, N_2 et n si et seulement si :

$$X(\Omega) = \{sup(0,\ n - N_2),\ \ldots,\ inf(n,\ N_1)\}$$

et

$$\forall k \in X(\Omega), \quad p(X = k) = \frac{C_{N_1}^{k}\, C_{N_2}^{n-k}}{C_{N_1+N_2}^{n}}$$

On note $X \to \mathcal{H}(N_1, N_2, n)$.

Proposition

Soit X une variable aléatoire qui suit une loi hypergéométrique de paramètres N_1, N_2 et n. Alors on a :

$$E(X) = np$$

$$V(X) = npq\frac{N-n}{N-1}$$

avec

$$N = N_1 + N_2, \quad p = \frac{N_1}{N}, \quad q = 1 - p$$

Exemple

Un joueur coche une grille de loto : il choisit 6 numéros parmi $\{1, 2, \ldots, 49\}$.
Parmi les 49 numéros, il y a 6 numéros gagnants (succès) et 43 numéros non gagnants.
1) Calculer la probabilité qu'a le joueur pour obtenir k numéros gagnants,
2) En moyenne, combien de numéros gagnants obtient-on en jouant une grille de loto ?

Notons X la variable aléatoire correspondant au nombre de numéros gagnants.
On a $X \to \mathcal{H}(6, 43, 6)$, Donc

$$X(\Omega) = \{0, \ldots, 6\}$$

et

$$p(X = k) = \frac{C_6^k C_{43}^{6-k}}{C_{49}^6}$$

On obtient :

k	0	1	2	3
$P(X=k)$	0.436	0.413	0.132	0.0177

k	4	5	6
$p(X=k)$	$9,69.10^{-4}$	$1,84.10^{-5}$	$7,15.10^{-8}$

On a $E(X) = np = 6\frac{6}{49} \simeq 0.735$
Donc en moyenne, on obtient moins d'un numéro gagnant par grille cochée.

Remarque

Si le choix de n éléments dans l'ensemble $E = S \cup \overline{S}$ se fait successivement et avec remise alors on a :

$$X(\Omega) = \{0,\ 1,\ \ldots,\ n\}$$

$$p(X = k) = C_n^k \frac{N_1^k (N_2)^{n-k}}{N^n}$$

On retrouve la loi binomiale.

VI- Couple de variables aléatoires

VI.1- Loi d'un couple aléatoire

Définition

Soit $(\Omega, \mathcal{T}, p)$ *un espace probabilisé et soient* X *et* Y *deux variables aléatoires discrètes sur* Ω.
On appelle loi de probabilité du couple (X, Y) *l'application, notée* $p_{(X,Y)}$, *définie de* $X(\Omega) \times Y(\Omega)$ *dans* $[0, 1]$ *par :*

$$\forall (x_i, y_j) \in X(\Omega) \times Y(\Omega), \quad p_{(X,Y)}(x_i, y_j) = p\left((X = x_i) \cap (Y = y_j)\right)$$

Les lois de probabilité de X et de Y s'appellent lois marginales du couple.

Remarques :

- Pour définir la loi de probabilité du couple (X, Y), il faut et il suffit de connaître les valeurs $p_{ij} = p_{(X,Y)}(x_i, y_j)$, $\forall\ (x_i, y_j) \in X(\Omega) \times Y(\Omega)$.
- On doit avoir $\displaystyle\sum_{i,j} p_{ij} = 1$.
- $\displaystyle p_{(X,Y)}(I \times J) = p((X \in I) \cap (Y \in J)) = \sum_{x_i \in I, y_j \in J} p_{ij}$
- Si $X(\Omega)$ et $Y(\Omega)$ sont finis, la loi du couple (X, Y) et les lois marginales de X et de Y peuvent être représentées par un tableau à deux dimensions appelé tableau de contingence :

$X \backslash Y$	y_1	y_2	$\dots$	y_j	$\dots$	y_m	Total
x_1	p_{11}	p_{12}	$\dots$	p_{1j}	$\dots$	p_{1m}	$p(x_1)$
x_2	p_{21}	p_{22}	$\dots$	p_{2j}	$\dots$	p_{2m}	$p(x_2)$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
x_i	p_{i1}	p_{i2}	$\dots$	p_{ij}	$\dots$	p_{im}	$p(x_i)$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
x_k	p_{k1}	p_{k2}	$\dots$	p_{kj}	$\dots$	p_{km}	$p(x_k)$
Total	$p(y_1)$	$p(y_2)$	$\dots$	$p(y_j)$	$\dots$	$p(y_m)$	1

On a :

$$p(x_i) = \sum_{j=1}^{m} p_{ij}$$

$$p(y_j) = \sum_{i=1}^{k} p_{ij}$$

$$\sum_{i=1}^{k} p(x_i) = \sum_{j=1}^{m} p(y_j) = \sum_{i=1}^{k} \sum_{j=1}^{m} p_{ij} = 1$$

Exemple

On place au hasard deux boules numérotés 1 et 2 dans deux boîtes A et B.
X désigne le nombre de boîtes vides.
Y désigne le nombre de billes dans la boîte A.

VI.2- Variables aléatoires indépendantes

Définition

Deux variables aléatoires X et Y sont dites indépendantes si et seulement si $\forall (x_i, y_j) \in X(\Omega) \times Y(\Omega)$, les événements $(X = x_i)$ et $(Y = y_j)$ sont indépendants.

Corollaire

Deux variables aléatoires X et Y sont dites indépendantes si et seulement si

$$\forall (x_i, y_j) \in X(\Omega) \times Y(\Omega), \quad p(x_i, y_j) = p(x_i) \times p(y_j)$$

On dit alors que la loi du couple (X, Y) est égale au produit des lois marginales.

VI.3- Covariance

Définition

On appelle covariance des variables aléatoires X et Y le nombre réel (si il existe), noté $COV(X, Y)$, défini par :

$$COV(X, Y) = E(XY) - E(X).E(Y)$$

On a :

$$E(XY) = \sum_{x_i \in X(\Omega),\ y_j \in Y(\Omega)} x_i.y_j.p(x_i, y_j)$$

Propriétés :

Soient X et Y deux variables aléatoires et a et b deux constantes réelles, on a :

- $COV(X + a, Y + b) = COV(X, Y)$
- $COV(aX, bY) = ab.COV(X, Y)$
- E(XY)=E(X).E(Y)+COV(X,Y)
- V(X+Y)=V(X)+V(Y)+2COV(X,Y)
- Si X et Y sont indépendantes alors $COV(X, Y) = 0$
  La réciproque est fausse.

Exemple

$X \backslash Y$	0	1	2	
0	0	$\frac{1}{2}$	0	$\frac{1}{2}$
1	$\frac{1}{4}$	0	$\frac{1}{4}$	$\frac{1}{2}$
	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$	1

On a : $E(X) = 1$, E(Y)=1/2, et E(XY)=1/2.
Cependant X et Y ne sont pas indépendantes.

Exercice

La loi du couple de variables (X, Y) est définie par le tableau suivant :

$X \backslash Y$	1	2	3
1	0	$\frac{1}{2}$	0
2	$\frac{1}{4}$	0	$\frac{1}{4}$

1) Déterminer les lois des variables X et Y.
2) X et Y sont-elles indépendantes ?
3) Calculer la covariance de X et Y. Conclure.

Exercice

Soient X et Y deux variables aléatoires indépendantes définies sur un même univers, dont les lois de probabilité sont données par :

x_i	0	1	2	3
$p(x_i)$	0.1	0.2	0.3	0.4

y_i	1	2	3
$P(y_i)$	0.25	0.5	0.25

1) Calculer l'espérance et la variance des deux variables.
2) Donner la loi de probabilité conjointe de X et Y.
3) Établir la loi de probabilité de $T = X + Y$. Calculer $E(T)$ et $V(T)$.
4) Établir la loi de probabilité de $Z = XY$. Calculer $E(Z)$ et $V(Z)$.
5) Calculer $Cov(X, Y)$.

Chapitre 5

Variables aléatoires continues

Introduction

Dans ce chapitre, on suppose que l'ensemble $X(\Omega)$ est infini non dénombrable.

Exemple

On lance une flèche sur une cible matérialisée par un disque de rayon R. On a :

$$\Omega = \{M(x, y) \ / \ x^2 + y^2 \leq R\}$$

On note X la variable aléatoire qui désigne la distance du point d'impact de la flèche au centre du disque. On a :

$$X(\Omega) =$$

Exemple

On lance une flèche sur une cible matérialisée par un disque de rayon R. On a :

$$\Omega = \{M(x, y) \ / \ x^2 + y^2 \leq R\}$$

On note X la variable aléatoire qui désigne la distance du point d'impact de la flèche au centre du disque. On a :

$$X(\Omega) = \ [0, R]$$

On dit que X est une variable aléatoire continue.

I- Définition

Définition

Soit Ω un univers infini non dénombrable.
Une variable aléatoire X définie sur Ω est dite continue si et seulement si l'ensemble $X(\Omega)$ est infini non dénombrable.

Exemple :

- On lance deux dés et on s'intéresse à la distance qui les sépare.

- On mesure la glycémie dans le sang d'un individu.

II- Densité de probabilité

Définition

Une fonction f est une densité de probabilité ssi

1. $f(t) \geq 0 \quad \forall t \in \mathbb{R}$,
2. $\displaystyle\int_{-\infty}^{+\infty} f(t)dt = 1.$

Définition

Soient f une densité de probabilité et X une variable aléatoire continue. On dit que f est une densité de X ssi

$$\forall\ (a, b) \in \mathbb{R}^2,\ \ p(a \leq X \leq b) = \int_a^b f(t)dt$$

Remarque

La densité de X définit complétement la loi de probabilité de X.

III- Fonction de répartition

Définition

Soit X une variable aléatoire continue de densité f.

On appelle fonction de répartition de X, l'application, notée F, définie de $\mathbb{R}$ vers $[0,1]$ par :

$$\forall\, x \in \mathbb{R}, \quad F(x) = p(X \leq x) = \int_{-\infty}^{x} f(t)dt$$

On a : $f(x) = F^{'}(x)$

Remarque

La fonction de répartition de X détermine complètement la loi de probabilité de X.

Propriétés :

Soit X une variable aléatoire continue, F sa fonction de répartition. On a :

- F est continue sur $\mathbb{R}$,
- F est croissante sur $\mathbb{R}$,
- $\lim\limits_{x \to -\infty} F(x) = 0$
- $\lim\limits_{x \to +\infty} F(x) = 1$,

Propriétés :

Soit X une variable aléatoire continue, F sa fonction de répartition et f sa densité de probabilité. On a :

- $p(X \leq a) = p(X < a) = F(a) = \int_{-\infty}^{a} f(t)dt,$
- $p(X \geq a) = p(X > a) = 1 - F(a) = \int_{a}^{+\infty} f(t)dt,$
- $p(a < X \leq b) = p(a \leq X \leq b) = p(a < X < b) = p(a \leq X < b) = F(b) - F(a) = \int_{a}^{b} f(t)dt,$
- $p(X = a) = 0,$

Exemple :

Soit X une variable aléatoire dont la densité de probabilité est donnée par :

$$f(x) = \begin{cases} ke^{-2x} & \text{si } x \geq 0 \\ 0 & \text{sinon} \end{cases}$$

1. Déterminer k et représenter $f(x)$.
2. Déterminer et représenter la fonction de répartition $F(x)$.
3. Calculer les probabilités des événements $(X \geq 2)$ et $(-1 \leq X < 2)$.

IV- Paramètres d'une variable aléatoire continue

IV.1- Espérance mathématique

Soit X une variable aléatoire continue et soit f sa densité de probabilité.

Définition

L'espérance mathématique de X (si elle existe) est définie par :

$$E(X) = \int_{-\infty}^{+\infty} t.f(t)dt$$

Si φ est une fonction réelle continue, alors

$$E(\varphi(X)) = \int_{-\infty}^{+\infty} \varphi(t)f(t)dt,$$

IV.2- Variance et écart-type

Définition

On suppose que l'espérance de X existe.
La variance de X (si elle existe) est définie par :

$$V(X) = E(X^2) - (E(X))^2$$

Définition

Si la variance de X existe, l'écart-type de X est défini par :

$$\sigma(X) = \sqrt{V(X)}$$

Propriétés

Soient X et Y deux variables aléatoires continues et a et b deux constantes réelles. On a :

- $E(aX + b) = aE(X) + b$,
- $V(aX + b) = a^2 V(X)$,
- $E(X + Y) = E(X) + E(Y)$,

Si X et Y sont indépendantes alors :

- $E(XY) = E(X)E(Y)$
- $V(X+Y) = V(X) + V(Y)$.

Exemple :

Calculer l'espérance et l'écart-type de la variable aléatoire de l'exemple précédent.

Exercice :

Soit X une variable aléatoire dont la densité de probabilité est définie par :

$$f(t) = \begin{cases} \frac{1}{k} & \text{si } t \in [1, 2[\\ \frac{1}{4} & \text{si } t \in [2, 4[\\ 0 & \text{sinon} \end{cases}$$

1) Déterminer la constante k.

2) Déterminer la fonction de répartition de X.

3) Calculer les probabilités suivantes :
$p(X \geq 3)$, $p(1 < X < 3)$, $p(3 \leq X \leq 5)$.

4) Calculer l'espérance et la variance de X.

Soit Y une variable aléatoire définie par $Y = 4 - X$.

5) Déterminer la fonction de répartition et la fonction densité de probabilité de Y.

6) Calculer l'espérance et la variance de Y.

V- Lois usuelles continues

V.1- Loi uniforme

Définition

Soit X une variable aléatoire continue. On dit que X suit la loi uniforme sur l'intervalle $[a, b]$ ssi sa densité de probabilité est définie par :

$$f_X(t) = \begin{cases} \dfrac{1}{b-a} & si\ t \in [a, b] \\ 0 & sinon \end{cases}$$

On note $X \to \mathcal{U}([a, b])$.

Proposition

Soit X une variable aléatoire continue qui suit une loi uniforme sur l'intervalle $[a, b]$. Alors on a :

$$F_X(t) = \begin{cases} 0 & si\ t \leq a \\ \dfrac{t-a}{b-a} & si\ t \in [a, b] \\ 1 & si\ t \geq b \end{cases}$$

$$E(X) = \frac{a+b}{2}$$

$$V(X) = \frac{(a-b)^2}{12}$$

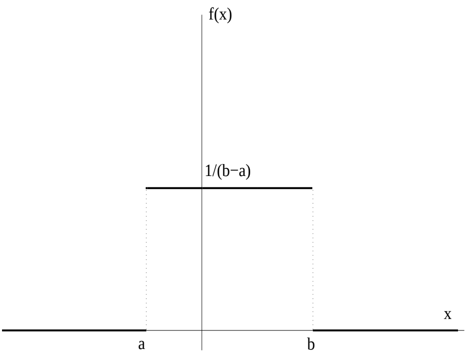


Figure: Densité de la loi uniforme.

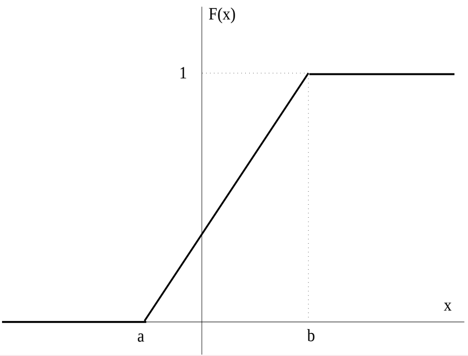


Figure: Fonction de répartition de la loi uniforme.

V.2- Loi exponentielle

Définition

On dit qu'une variable aléatoire continue X suit la loi exponentielle de paramètre λ ssi sa densité de probabilité est définie par :

$$f_X(t) = \begin{cases} \lambda e^{-\lambda t} & si\ t \geq 0 \\ 0 & sinon \end{cases}$$

On note $X \rightarrow \mathcal{E}(\lambda)$.

Proposition

Soit X une variable aléatoire continue qui suit une loi exponentielle de paramètre λ. Alors on a :

$$F_X(t) = \begin{cases} 1 - e^{-\lambda t} & si\ t \geq 0 \\ 0 & sinon \end{cases}$$

$$E(X) = \frac{1}{\lambda}$$

$$V(X) = \frac{1}{\lambda^2}$$

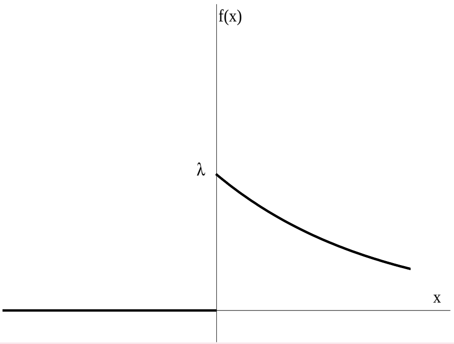


Figure: Densité de la loi exponentielle.

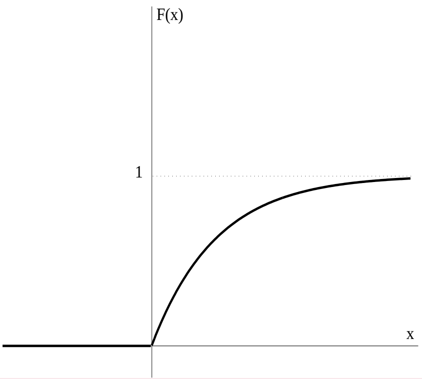


Figure: Fonction de répartition de la loi exponentielle.

V.3- Loi Normale ou loi de Gauss

Définition

On dit qu'une variable aléatoire continue X suit la loi normale de paramètres m et σ ssi sa densité de probabilité est définie par :

$$f_X(t) = \frac{1}{\sigma\sqrt{2\pi}} exp\left(-\frac{(t-m)^2}{2\sigma^2}\right) \quad , \quad \forall\ t\ \in \mathbb{R}$$

On note $X \rightarrow \mathcal{N}(m, \sigma)$.

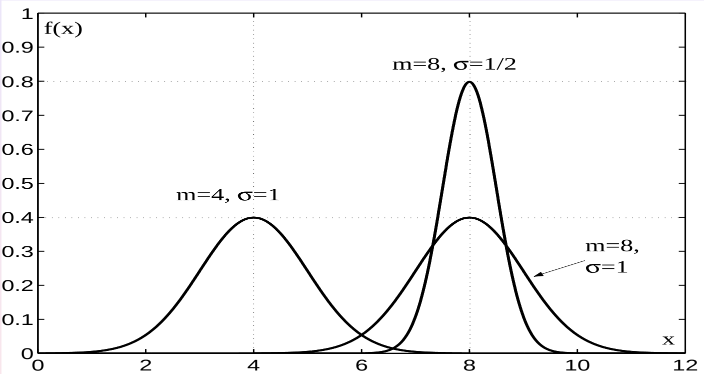


Figure: Densité de probabilité de la loi normale pour différentes valeurs de m et σ

Proposition

Soit X une variable aléatoire continue qui suit une loi normale de paramètres m et σ. Alors on a :

$$E(X) = m$$

$$V(X) = \sigma^2$$

Remarque

Soit X une variable aléatoire continue qui suit une loi normale de paramètres m et σ. Alors on a :

$$F_X(t) = \int_{-\infty}^{t} \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(u-m)^2}{2\sigma^2}\right) du$$

La fonction de répartition d'une loi normale $\mathcal{N}(m, \sigma)$ ne peut pas être exprimée à l'aide de fonctions usuelles.

Remarque

Pour surmonter cette difficulté, nous utilisons des tables numériques donnant les valeurs de la fonction de répartition d'une loi normale centrée réduite $\mathcal{N}(0, 1)$.

Pour calculer numériquement les valeurs de la fonction de répartition d'une loi normale $\mathcal{N}(m, \sigma)$ quelconques, nous utilisons les propositions suivantes :

Proposition

Si $X \to \mathcal{N}(m, \sigma)$ alors la variable aléatoire centrée réduite $X^* = \dfrac{X - m}{\sigma}$ suit la loi normale $\mathcal{N}(0, 1)$.

$\mathcal{N}(0, 1)$ est appelée loi normale centrée réduite.

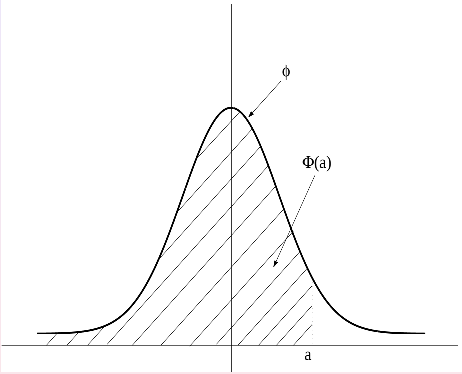


Figure: Densité et fonction de répartition de la loi normale $\mathcal{N}(0, 1)$

- On note $\varphi(t)$ et $\phi(t)$, respectivement, la densité et la fonction de répartition de la loi $\mathcal{N}(0,1)$. on a :

$$\varphi(t) = \frac{1}{\sqrt{2\pi}} exp\left(-\frac{t^2}{2}\right)$$

$$\phi(t) = \int_{-\infty}^{t} \varphi(u)du$$

- On dispose de tables numériques donnant les valeurs $\varphi(t)$ et $\phi(t)$ pour t variant de 0 à 5.9

- Pour les valeurs négatives de t, on utilise la proposition suivante :

Proposition

$$\varphi(-t) = \varphi(t)$$

$$\phi(-t) = 1 - \phi(t)$$

Proposition

Soit $X \to \mathcal{N}(m, \sigma)$. Alors les valeurs de la fonction de répartition $F_X(t)$ et de la fonction densité de probabilité $f_X(t)$ de X se déduisent de celles de la loi normale centrée réduite $\mathcal{N}(0, 1)$ à l'aide des relations suivantes :

$$F_X(t) = \phi\left(\frac{t-m}{\sigma}\right)$$

$$f_X(t) = \frac{1}{\sigma}\,\varphi\left(\frac{t-m}{\sigma}\right)$$

Utilisation de la table numérique de $\phi(t)$

- Pour les valeurs de t variant de 0 à 2.99, on lit les décimales dans les lignes de la première colonne et les centièmes dans les colonnes de la première ligne.

Utilisation de la table numérique de $\phi(t)$

- Pour les valeurs de t variant de 0 à 2.99, on lit les décimales dans les lignes de la première colonne et les centièmes dans les colonnes de la première ligne.

  Par exemple, la valeur de $\phi(1.34)$ se trouve à l'intersection de la ligne 1.3 et de la colonne 0.04

Utilisation de la table numérique de $\phi(t)$

- Pour les valeurs de t variant de 0 à 2.99, on lit les décimales dans les lignes de la première colonne et les centièmes dans les colonnes de la première ligne.

  Par exemple, la valeur de $\phi(1.34)$ se trouve à l'intersection de la ligne 1.3 et de la colonne 0.04

  On trouve $\phi(1.34) = 0.90988$

- Pour les valeurs de t comprises entre 3 et 5.9, la table donne $1 - \phi(t)$. On lit la partie entière dans les lignes de la première colonne et les décimales dans les colonnes de la première ligne. La notation $\{m; k\}$ signifie $m\,10^{-k}$.

- Pour les valeurs de t comprises entre 3 et 5.9, la table donne $1 - \phi(t)$. On lit la partie entière dans les lignes de la première colonne et les décimales dans les colonnes de la première ligne. La notation $\{m; k\}$ signifie $m\,10^{-k}$.

  Par exemple, la valeur de $1 - \phi(4.6)$ se trouve à l'intersection de la ligne 4 et de la colonne 0.6

- Pour les valeurs de t comprises entre 3 et 5.9, la table donne $1 - \phi(t)$. On lit la partie entière dans les lignes de la première colonne et les décimales dans les colonnes de la première ligne. La notation $\{m; k\}$ signifie $m\ 10^{-k}$.

  Par exemple, la valeur de $1 - \phi(4.6)$ se trouve à l'intersection de la ligne 4 et de la colonne 0.6

  On trouve $\{2.11; 6\}$, c'est à dire $\phi(4.6) = 1 - 2.11\ 10^{-6}$.

- Pour les valeurs de t comprises entre 0 et 5.9 avec plus de 2 chiffres après la virgule, on détermine $\phi(t)$ soit par arrondi soit par interpolation.

- Pour les valeurs de t comprises entre 0 et 5.9 avec plus de 2 chiffres après la virgule, on détermine $\phi(t)$ soit par arrondi soit par interpolation.

  Par exemple, la valeur de $\phi(1.3476)$ peut être soit approximée par $\phi(1.35)$ (on trouve 0.91149), soit calculée par interpolation à partir des valeurs de $\phi(1.34)$ et $\phi(1.35)$ (on trouve 0.91110).

- Pour les valeurs de t comprises entre 0 et 5.9 avec plus de 2 chiffres après la virgule, on détermine $\phi(t)$ soit par arrondi soit par interpolation.

  Par exemple, la valeur de $\phi(1.3476)$ peut être soit approximée par $\phi(1.35)$ (on trouve 0.91149), soit calculée par interpolation à partir des valeurs de $\phi(1.34)$ et $\phi(1.35)$ (on trouve 0.91110).
- Pour les valeurs négatives, on utilise $\phi(-t) = 1 - \phi(t)$.

- Pour les valeurs de t comprises entre 0 et 5.9 avec plus de 2 chiffres après la virgule, on détermine $\phi(t)$ soit par arrondi soit par interpolation.

  Par exemple, la valeur de $\phi(1.3476)$ peut être soit approximée par $\phi(1.35)$ (on trouve 0.91149), soit calculée par interpolation à partir des valeurs de $\phi(1.34)$ et $\phi(1.35)$ (on trouve 0.91110).
- Pour les valeurs négatives, on utilise $\phi(-t) = 1 - \phi(t)$.
- Pour les valeurs de $t > 5.9$, on considère que $\phi(t) \simeq 1$.

Utilisation de la table numérique de $\varphi(t)$

- Pour les valeurs de t comprises entre 0 et 2.99, on lit les décimales dans les lignes de la première colonne et les centièmes dans les colonnes de la première ligne.

Utilisation de la table numérique de $\varphi(t)$

- Pour les valeurs de t comprises entre 0 et 2.99, on lit les décimales dans les lignes de la première colonne et les centièmes dans les colonnes de la première ligne.

  Par exemple, la valeur de $\varphi(1.34)$ se trouve à l'intersection de la ligne 1.3 et de la colonne 0.04

Utilisation de la table numérique de $\varphi(t)$

- Pour les valeurs de t comprises entre 0 et 2.99, on lit les décimales dans les lignes de la première colonne et les centièmes dans les colonnes de la première ligne.

  Par exemple, la valeur de $\varphi(1.34)$ se trouve à l'intersection de la ligne 1.3 et de la colonne 0.04

  On trouve $\varphi(1.34) = 0.16256$

- Pour les valeurs de t comprises entre 3 et 5.9, on lit la partie entière dans les lignes de la première colonne et les décimales dans les colonnes de la première ligne. La notation $\{m; k\}$ signifie $m\,10^{-k}$.

- Pour les valeurs de t comprises entre 3 et 5.9, on lit la partie entière dans les lignes de la première colonne et les décimales dans les colonnes de la première ligne. La notation $\{m; k\}$ signifie $m\,10^{-k}$.

  Par exemple, la valeur de $\varphi(4.6)$ se trouve à l'intersection de la ligne 4 et de la colonne 0.6

- Pour les valeurs de t comprises entre 3 et 5.9, on lit la partie entière dans les lignes de la première colonne et les décimales dans les colonnes de la première ligne. La notation $\{m; k\}$ signifie $m\ 10^{-k}$.

  Par exemple, la valeur de $\varphi(4.6)$ se trouve à l'intersection de la ligne 4 et de la colonne 0.6

  On trouve $\{1.01; 5\}$, c'est à dire $\varphi(4.6) = 1.01\ 10^{-5}$.

- Pour les valeurs de t comprises entre 0 et 5.9 avec plus de 2 chiffres après la virgule, on détermine $\varphi(t)$ soit par arrondi soit par interpolation.

- Pour les valeurs de t comprises entre 0 et 5.9 avec plus de 2 chiffres après la virgule, on détermine $\varphi(t)$ soit par arrondi soit par interpolation.
- Pour les valeurs de $t > 5.9$, on considère que $\varphi(t) \simeq 0$.

- Pour les valeurs de t comprises entre 0 et 5.9 avec plus de 2 chiffres après la virgule, on détermine $\varphi(t)$ soit par arrondi soit par interpolation.
- Pour les valeurs de $t > 5.9$, on considère que $\varphi(t) \simeq 0$.
- Pour les valeurs négatives, on utilise $\varphi(-t) = \varphi(t)$.

Exemples

- Soit $X \to \mathcal{N}(3.5, 2)$. Calculer $P(X < 4)$.

- Soit $X \to \mathcal{N}(50, 4)$. Calculer $P(40 < X < 60)$.

VI- Approximations par une loi normale

VI.1- Théorème de la limite centrale

Théorème

Soit $X_1, X_2, \ldots, X_n$ n variables aléatoires indépendantes, de même loi de probabilité, donc de même espérance m et de même variance σ^2.

on considère la variable aléatoire $S_n = \sum_{i=1}^{n} X_i$. Alors :

Pour n assez grand, la variable aléatoire S_n converge vers la loi $\mathcal{N}(nm, \sqrt{n}\sigma)$.

On note $S_n \approx \mathcal{N}(nm, \sqrt{n}\sigma)$.

Exemple

Un caissier constate une erreur d'au plus 10 dh par jours (en plus ou en moins).
Quelle est la probabilité pour que l'erreur commise au bout de 220 jours soit inférieure à 200 dh ?

VI.2- Approximation de la loi binomiale

Proposition

Soit $X \to \mathcal{B}(n, p)$. Alors :
Pour n suffisamment grand ($n > 50$) et p et q pas trop proches de zéro ($np > 5$ et $nq > 5$), la loi binomiale $\mathcal{B}(n, p)$ peut être approximée par la loi normale de paramètres $m = np$ et $\sigma = \sqrt{npq}$.

$$X \approx \mathcal{N}(np, \sqrt{npq}).$$

$$p(X = k) \simeq \frac{1}{\sigma}\,\varphi\left(\frac{k-m}{\sigma}\right), \quad 0 \leq k \leq n$$

$$p(k_1 \leq X \leq k_2) \simeq \phi\left(\frac{k_2-m}{\sigma}\right) - \phi\left(\frac{k_1-m}{\sigma}\right), \quad 0 \leq k_1 < k_2 \leq n$$

Exemple

On lance un dé 4500 fois. On appelle X la variable aléatoire qui compte le nombre d'apparition de la face N° 1.

1) Quelle est la loi de probabilité de X ?

2) Calculer les probabilités suivantes :
- $p(X = 1000)$ - $p(980 \leq X \leq 1030)$,
- $p(980 < X < 1030)$.

VI.3- Approximation de la loi Poisson

Proposition

Soit $X \to \mathcal{P}(\lambda)$. Alors :
Pour λ suffisamment grand ($\lambda > 20$), la loi de Poisson $\mathcal{P}(\lambda)$ peut être approximée par la loi normale de paramètres $m = \lambda$ et $\sigma = \sqrt{\lambda}$.

$X \approx \mathcal{N}(\lambda, \sqrt{\lambda})$.

$$p(X = k) \simeq \frac{1}{\sigma} \varphi \left(\frac{k - m}{\sigma} \right), \quad k \in \mathbb{N}$$

$$p(k_1 \leq X \leq k_2) \simeq \phi \left(\frac{k_2 - m}{\sigma} \right) - \phi \left(\frac{k_1 - m}{\sigma} \right), \quad k_1 < k_2$$

Exemple

Une clinique traite en moyenne deux urgences par jour. Quelle est la probabilité pour que la clinique traite plus de 70 urgences par mois ?

C.1 Loi normale

Fonction de répartition de la loi normale réduite : $\Phi(x) = \dfrac{1}{\sqrt{2\pi}} \displaystyle\int_{-\infty}^{x} \exp\Big(-\frac{t^2}{2}\Big)\, dt$

x	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.	0.5	0.50399	0.50798	0.51197	0.51595	0.51994	0.52392	0.52790	0.53188	0.53586
0.1	0.53983	0.54380	0.54776	0.55172	0.55567	0.55962	0.56356	0.56749	0.57142	0.57535
0.2	0.57926	0.58317	0.58706	0.59095	0.59483	0.59871	0.60257	0.60642	0.61026	0.61409
0.3	0.61791	0.62172	0.62552	0.62930	0.63307	0.63683	0.64058	0.64431	0.64803	0.65173
0.4	0.65542	0.65910	0.66276	0.66640	0.67003	0.67364	0.67724	0.68082	0.68439	0.68793
0.5	0.69146	0.69497	0.69847	0.70194	0.70540	0.70884	0.71226	0.71566	0.71904	0.72240
0.6	0.72575	0.72907	0.73237	0.73565	0.73891	0.74215	0.74537	0.74857	0.75175	0.75490
0.7	0.75804	0.76115	0.76424	0.76730	0.77035	0.77337	0.77637	0.77935	0.78230	0.78524
0.8	0.78814	0.79103	0.79389	0.79673	0.79955	0.80234	0.80511	0.80785	0.81057	0.81327
0.9	0.81594	0.81859	0.82121	0.82381	0.82639	0.82894	0.83147	0.83398	0.83646	0.83891
1.	0.84134	0.84375	0.84614	0.84849	0.85083	0.85314	0.85543	0.85769	0.85993	0.86214
1.1	0.86433	0.86650	0.86864	0.87076	0.87286	0.87493	0.87698	0.87900	0.88100	0.88298
1.2	0.88493	0.88686	0.88877	0.89065	0.89251	0.89435	0.89617	0.89796	0.89973	0.90147
1.3	0.90320	0.90490	0.90658	0.90824	0.90988	0.91149	0.91309	0.91466	0.91621	0.91774
1.4	0.91924	0.92073	0.92220	0.92364	0.92507	0.92647	0.92785	0.92922	0.93056	0.93189
1.5	0.93319	0.93448	0.93574	0.93699	0.93822	0.93943	0.94062	0.94179	0.94295	0.94408
1.6	0.94520	0.94630	0.94738	0.94845	0.94950	0.95053	0.95154	0.95254	0.95352	0.95449
1.7	0.95543	0.95637	0.95728	0.95818	0.95907	0.95994	0.96080	0.96164	0.96246	0.96327
1.8	0.96407	0.96485	0.96562	0.96638	0.96712	0.96784	0.96856	0.96926	0.96995	0.97062
1.9	0.97128	0.97193	0.97257	0.97320	0.97381	0.97441	0.97500	0.97558	0.97615	0.97670
2.	0.97725	0.97778	0.97831	0.97882	0.97932	0.97982	0.98030	0.98077	0.98124	0.98169
2.1	0.98214	0.98257	0.98300	0.98341	0.98382	0.98422	0.98461	0.98500	0.98537	0.98574
2.2	0.98610	0.98645	0.98679	0.98713	0.98745	0.98778	0.98809	0.98840	0.98870	0.98899
2.3	0.98928	0.98956	0.98983	0.99010	0.99036	0.99061	0.99086	0.99111	0.99134	0.99158
2.4	0.99180	0.99202	0.99224	0.99245	0.99266	0.99286	0.99305	0.99324	0.99343	0.99361
2.5	0.99379	0.99396	0.99413	0.99430	0.99446	0.99461	0.99477	0.99492	0.99506	0.99520
2.6	0.99534	0.99547	0.99560	0.99573	0.99585	0.99598	0.99609	0.99621	0.99632	0.99643
2.7	0.99653	0.99664	0.99674	0.99683	0.99693	0.99702	0.99711	0.99720	0.99728	0.99736
2.8	0.99744	0.99752	0.99760	0.99767	0.99774	0.99781	0.99788	0.99795	0.99801	0.99807
2.9	0.99813	0.99819	0.99825	0.99831	0.99836	0.99841	0.99846	0.99851	0.99856	0.99861

Table de $1 - \Phi(x)$ pour les grandes valeurs de x

x	0.	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
3.	{1.35 ; 3}	{9.68 ; 4}	{6.87 ; 4}	{4.83 ; 4}	{3.37 ; 4}	{2.33 ; 4}	{1.59 ; 4}	{1.08 ; 4}	{7.23 ; 5}	{4.81 ; 5}
4.	{3.17 ; 5}	{2.07 ; 5}	{1.33 ; 5}	{8.54 ; 6}	{5.41 ; 6}	{3.40 ; 6}	{2.11 ; 6}	{1.30 ; 6}	{7.93 ; 7}	{4.79 ; 7}
5.	{2.87 ; 7}	{1.70 ; 7}	{9.96 ; 8}	{5.79 ; 8}	{3.33 ; 8}	{1.90 ; 8}	{1.07 ; 8}	{5.99 ; 9}	{3.32 ; 9}	{1.82 ; 9}

N.B. La notation {m ;k} signifie $m\, 10^{-k}$.

Densité de la loi normale réduite : $\varphi(x) = \dfrac{1}{\sqrt{2\pi}} \exp\Big(-\dfrac{t^2}{2}\Big)$

x	0.	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.	0.39894	0.39892	0.39886	0.39876	0.39862	0.39844	0.39822	0.39797	0.39767	0.39733
0.1	0.39695	0.39654	0.39608	0.39559	0.39505	0.39448	0.39387	0.39322	0.39253	0.39181
0.2	0.39104	0.39024	0.38940	0.38853	0.38762	0.38667	0.38568	0.38466	0.38361	0.38251
0.3	0.38139	0.38023	0.37903	0.37780	0.37654	0.37524	0.37391	0.37255	0.37115	0.36973
0.4	0.36827	0.36678	0.36526	0.36371	0.36213	0.36053	0.35889	0.35723	0.35553	0.35381
0.5	0.35207	0.35029	0.34849	0.34667	0.34482	0.34294	0.34105	0.33912	0.33718	0.33521
0.6	0.33322	0.33121	0.32918	0.32713	0.32506	0.32297	0.32086	0.31874	0.31659	0.31443
0.7	0.31225	0.31006	0.30785	0.30563	0.30339	0.30114	0.29887	0.29659	0.29431	0.29200
0.8	0.28969	0.28737	0.28504	0.28269	0.28034	0.27798	0.27562	0.27324	0.27086	0.26848
0.9	0.26609	0.26369	0.26129	0.25888	0.25647	0.25406	0.25164	0.24923	0.24681	0.24439
1.	0.24197	0.23955	0.23713	0.23471	0.23230	0.22988	0.22747	0.22506	0.22265	0.22025
1.1	0.21785	0.21546	0.21307	0.21069	0.20831	0.20594	0.20357	0.20121	0.19886	0.19652
1.2	0.19419	0.19186	0.18954	0.18724	0.18494	0.18265	0.18037	0.17810	0.17585	0.17360
1.3	0.17137	0.16915	0.16694	0.16474	0.16256	0.16038	0.15822	0.15608	0.15395	0.15183
1.4	0.14973	0.14764	0.14556	0.14350	0.14146	0.13943	0.13742	0.13542	0.13344	0.13147
1.5	0.12952	0.12758	0.12566	0.12376	0.12188	0.12001	0.11816	0.11632	0.11450	0.11270
1.6	0.11092	0.10915	0.10741	0.10567	0.10396	0.10226	0.10059	0.09893	0.09728	0.09566
1.7	0.09405	0.09246	0.09089	0.08933	0.08780	0.08628	0.08478	0.08329	0.08183	0.08038
1.8	0.07895	0.07754	0.07614	0.07477	0.07341	0.07206	0.07074	0.06943	0.06814	0.06687
1.9	0.06562	0.06438	0.06316	0.06195	0.06077	0.05959	0.05844	0.05730	0.05618	0.05508
2.	0.05399	0.05292	0.05186	0.05082	0.04980	0.04879	0.04780	0.04682	0.04586	0.04491
2.1	0.04398	0.04307	0.04217	0.04128	0.04041	0.03955	0.03871	0.03788	0.03706	0.03626
2.2	0.03547	0.03470	0.03394	0.03319	0.03246	0.03174	0.03103	0.03034	0.02965	0.02898
2.3	0.02833	0.02768	0.02705	0.02643	0.02582	0.02522	0.02463	0.02406	0.02349	0.02294
2.4	0.02239	0.02186	0.02134	0.02083	0.02033	0.01984	0.01936	0.01888	0.01842	0.01797
2.5	0.01753	0.01709	0.01667	0.01625	0.01585	0.01545	0.01506	0.01468	0.01431	0.01394
2.6	0.01358	0.01323	0.01289	0.01256	0.01223	0.01191	0.01160	0.01130	0.01100	0.01071
2.7	0.01042	0.01014	0.00987	0.00961	0.00935	0.00909	0.00885	0.00861	0.00837	0.00814
2.8	0.00792	0.00770	0.00748	0.00727	0.00707	0.00687	0.00668	0.00649	0.00631	0.00613
2.9	0.00595	0.00578	0.00562	0.00545	0.00530	0.00514	0.00499	0.00485	0.00470	0.00457

Table de $\varphi(x)$ pour les grandes valeurs de x

x	0.	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
3.	{4.43 ; 3}	{3.27 ; 3}	{2.38 ; 3}	{1.72 ; 3}	{1.23 ; 3}	{8.73 ; 4}	{6.12 ; 4}	{4.25 ; 4}	{2.92 ; 4}	{1.99 ; 4}
4.	{1.34 ; 4}	{8.93 ; 5}	{5.89 ; 5}	{3.85 ; 5}	{2.49 ; 5}	{1.60 ; 5}	{1.01 ; 5}	{6.37 ; 6}	{3.96 ; 6}	{2.44 ; 6}
5.	{1.49 ; 6}	{8.97 ; 7}	{5.36 ; 7}	{3.17 ; 7}	{1.86 ; 7}	{1.08 ; 7}	{6.18 ; 8}	{3.51 ; 8}	{1.98 ; 8}	{1.10 ; 8}

N.B. La notation {m ;k} signifie $m\,10^{-k}$.